# Binding protein-dependent secondary transport in *Escherichia coli*: Quest for an elusive substrate

Titia H. Plantinga

*Foar Beppe*

*& foar myn âlders*

Cover: *Quest*

Cover design: Titia Plantinga

This work was performed in the Molecular Microbiology Group of the Groningen Biomolecular Sciences and Biotechnology Institute (GBB) of the University of Groningen, The Netherlands, and supported by the Netherlands Organisation for Scientific Research (NWO).

**RIJKSUNIVERSITEIT GRONINGEN**

# Binding protein-dependent secondary transport in *Escherichia coli*: Quest for an elusive substrate

**Proefschrift**

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
vrijdag 19 december 2003
om 16.00 uur

door

Tietje Hendrikje Plantinga

geboren op 27 maart 1974
te Het Bildt

Promotores: Prof. Dr. A.J.M. Driessen
Prof. Dr. W.N. Konings

Beoordelingscommissie: Prof. Dr. L. Dijkhuizen
Prof. Dr. O.P. Kuipers
Prof. Dr. M. Veenhuis

ISBN-nummer: 90-9017354-4

## *Voorwoord*

## *Preface*

*"Ahm... help, I need a little help here!"* [229]

Een lange tocht door bergachtig gebied. De top is reeds aan het begin van de tocht in zicht, maar schijnbaar onbereikbaar. Een pad dat vaker bergafwaarts dan bergopwaarts lijkt te gaan. Kortere routes, maar niet het juiste materiaal om de klim te kunnen maken. En er is geen routebeschrijving.

Kortom: een echte uitdaging. Gelukkig ben ik onderweg veel personen tegengekomen die me de weg wilden wijzen als dat nodig was. Ook al werd soms onbedoeld niet in de juiste richting gewezen.

Zo heb ik de top gehaald: zelfstandig, maar niet alleen.

Dus: iedereen die op wat voor manier dan ook een bijdrage heeft geleverd aan de totstandkoming van dit proefschrift: ik dank je hartelijk!

A long journey through rough terrain. Trying to reach the top of a mountain that is visible from the start, yet seemingly out of reach. The road winding downhill, it seems, more often than uphill. Shortcuts that can't be taken for lack of the right climbing gear. And there is no road map.

Faced with such a challenge, it has been comforting to know there were always people ready to give directions when I needed them. Even if sometimes unwillingly pointing me the wrong way.

This is how I made it to the top: independently, but not alone.

So, to everybody who in any way helped me finish this thesis: I thank you most sincerely!

Wees niet bevreesd langzaam vooruit te komen,
maar wees bevreesd stil te blijven staan [2]

# *Contents*

*"There is no such thing as a moral or an immoral book.*
*Books are well written, or badly written. That is all."* [234]

***Chapter 1***

# *Chapter 1: General Introduction*

## Binding protein-dependent transport systems: Structure, mechanism, and evolution

*"What is a man, if his chief good and market of his time*
*Be but to sleep and feed? a beast, no more*
*Sure he that made us with such large discourse*
*Looking before and after, gave us not*
*That capability and god-like reason*
*To fust in us unus'd."* [189]

## *Contents*

## 1. An introduction to prokaryotic binding protein-dependent transporters

Prokaryotes depend on cell surface-localized transport proteins for the uptake of nutrients and other factors that are vital for survival, as well as for the export of metabolic endproducts and potentially toxic compound [54, 124]. In a number of cases bidirectional transport is mediated by a single membrane protein. The ability to transport a broad range of substrates provides the cell with a selective advantage in the ongoing struggle for survival. The genes encoding transport systems make up a significant part of prokaryotic genomes [159]. Historically transport proteins have been classified according to the manner in which transport is energized, as well as the architecture of the structural components. To date, four major classes of prokaryotic solute uptake transporters have been identified (Fig. 1). First, secondary transporters consist of a single permease domain, and transport is driven by the proton (pmf) and/or sodium ion (smf) motive force (Fig. 1a) [157, 163]. Second, binding protein-dependent secondary transporters energize transport in a similar manner, but the substrate is delivered to the integral membrane permease by a high-affinity substrate-binding protein (Fig. 1b) [53, 98, 171]. Third, binding protein-dependent primary transporters also utilize a substrate-binding protein, but transport is energized by ATP-hydrolysis (Fig. 1c). These proteins are members of the ATP-binding cassette (ABC) superfamily of transport proteins [7, 40, 77, 78, 86]. Fourth and finally, the bacterial phospho*enol*pyruvate:carbohydrate phosphotransferase systems (PTS) transfer a phosphoryl group from phospho*enol*pyruvate (PEP) via several intermediate proteins to their carbohydrate substrates during uptake (Fig. 1d) [164, 165, 181]. The latter transport systems will not be discussed any further, as these are not part of the scope of this thesis.

The binding protein-dependent secondary transporters share a common structural organisation (Fig. 1b), while increasing information concerning their functional diversity gradually becomes available [53, 54, 98, 160, 171]. Likewise, binding protein-dependent ABC transporters display a remarkable similarity in their general domain composition (Fig. 1c), whereas the potential substrates of these transporters are highly diverse [77, 78, 170, 200].

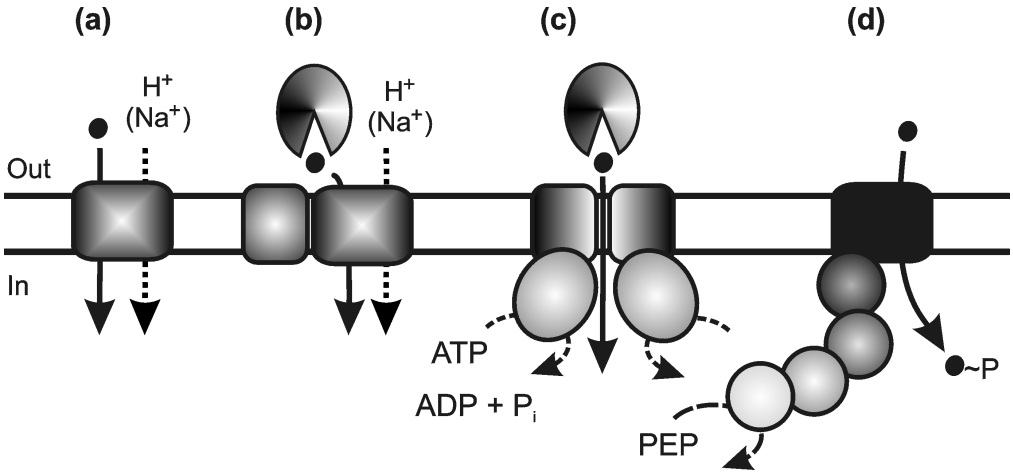



**Figure 1.** Four classes of prokaryotic solute transport systems (see text for details). (a) Secondary transporter. (b) Binding protein-dependent secondary transporter. (c) Binding protein-dependent primary transporters. (d) Phospho*enol*pyruvate:carbohydrate phosphotransferase systems. During transport the phosphate from phospho*enol*puruvate (PEP) is transferred, via several intermediate proteins, to the carbohydrate substrate.

The respective structural components of the transporter supply the important characteristics of these systems. The high affinity is ensured by an extramembranal substrate-binding protein [147, 170, 200], which delivers the substrate to the permease that facilitates the actual transmembrane transport reaction [54, 185, 186]. The energetic requirements for the transport process are met by the nucleotide-binding domains (NBDs) of the ABC transporters [42, 77], or by coupling to the pmf and/or smf which probably takes place through the large subunit of the binding protein-dependent secondary transporters (see section 3.4) [53, 54, 171].

### *1.1 Extramembranal substrate-binding domains*

Two of the prokaryotic transport protein classes make use of a periplasmic (Gram-negative bacteria) or extracellular (Gram-positive bacteria and archaea) high-affinity substrate-binding protein (Fig. 1b & c) [7, 54, 77, 98, 170, 171]. Induction of expression of these binding protein-dependent systems often occurs in the absence of the transported substrate, suggesting that these systems are used for scavenging compounds that are present at low concentrations [147]. The extramembranal

substrate-binding proteins of bacteria and archaea have been linked to functions other than substrate transport alone. Some have been implicated in chemotaxis and/or sensory signal transduction [200], whereas others show chaperone-like properties by interacting with periplasmic proteins that have not folded correctly [174]. The focus here, however, remains on the solute transport function of this class of proteins.

In Gram-negative bacteria, the substrate-binding proteins are produced with an N-terminal signal sequence for secretion into the periplasm that is removed during translocation via the Sec-system [130]. The periplasm of Gram-negative bacteria is a gel-like, continuous compartment that allows only slow lateral diffusion of its components. The diffusion rate of the maltose-binding protein MalE (MBP) of *Escherichia coli* in this compartment is $10^3$ times slower than in water, and $10^2$ times slower than in the cytoplasm [20, 29, 82]. Thus, the substrate-binding proteins of these bacteria are moving more or less freely, yet relatively slowly through the periplasm [7, 200]. Efficient transport by these systems would require a high concentration of substrate-binding proteins at the permease. This is in part achieved by the high affinity of these proteins for the permease, even in the absence of substrate, as was shown for the histidine-binding protein HisJ of *Salmonella typhimurium* [9]. As the solute can traverse more easily through the periplasm than the much larger substrate-binding domain, a higher efficiency of solute binding and subsequent transfer to the permease domains may be achieved in this way. In addition, the substrate-binding proteins are expressed in excess compared to the membrane-localized permeases to ensure high efficiency of solute transport [9].

In Gram-positive bacteria, which do not have an outer membrane, the substrate-binding proteins are usually tethered to the cell membrane via a lipid anchor to prevent loss of protein [77, 199, 200]. Interestingly, a completely different structure is found for the glycine betaine transporter OpuA from *Lactococcus lactis*: here the substrate-binding domain (OpuAB) is C-terminally fused to the membrane domain (OpuAC) [212]. This member of the OTCN (osmoprotectants, taurine, cyanate, nitrate) subfamily of the ABC superfamily [40] was recently shown not to be unique: a genome survey has shown that their occurrence is widespread [153, 213]. In addition, several proteins were found where even more than one solute-binding subunit was fused to a single membrane domain [213]. In archaea, binding proteins appear to be membrane connected via a

transmembrane segment [5] although lipid anchoring seems also widespread in these organisms. As in the generally accepted view on ABC transporters the membrane domains are only functional permeases as homo- or heterodimers, in these systems multiple substrate-binding domains will be active per permease.

*1.2 Questions in prokaryotic solute transport*

The structural composition of binding protein-dependent transporters raises questions about the mechanism of transport. These concern the mechanism of high-affinity binding of substrate and its subsequent delivery to the permease by the substrate-binding domain, as well as the translocation of the solute across the membrane. In addition, the mode of energization of the complete transport event is a major focus of research. However, purified binding protein-dependent transporters have proven to be unwilling subjects for biochemical studies, due to their multi-subunit composition [4, 49], and only a few reports exist that make use of a purified and membrane-reconstituted transporter complex. The best studied multicomponent solute uptake systems known to date are the histidine transporter HisJQMP of *S. typhimurium* and the maltose transporters MalEFGK of *E. coli* and *S. typhimurium*. These were the first to be succesfully overexpressed, purified and reconstituted into artificial membranes allowing for detailed studies concerning transport activity [47, 111, 118]. These and other studies will be discussed in more detail below. A recent major development is the availability of structural information on both the membrane-embedded permease and the solute-binding protein of the vitamin $B_{12}$ transporter BtuCDF of *E. coli* [27, 97]. These data supply molecular information on the interaction of these domains during the transport process.

In addition, mechanistic information obtained from work on structurally related permeases that do not require a solute-binding protein may be used to address the questions raised above, since their mechanism may follow general rules. As the body of data obtained from experiments ranging from genomic surveys to high-resolution crystal structures of integral membrane permeases increases, slowly but certainly the mechanistic details of these intriguing transporters will be unveiled.

This chapter compiles the information that is available to date on prokaryotic binding protein-dependent transport proteins. First, three of the major classes of prokaryotic solute uptake systems are described in further detail. Subsequently, the (proposed) mechanisms of transport by prokaryotic binding protein-dependent uptake systems are discussed. Finally, we will speculate on the evolution of these different classes of transport proteins known to date.

## 2. Secondary transporters

Secondary transporters are present in all three kingdoms of life. These systems are able to accumulate solutes in a pmf and/or smf-dependent manner [129, 157, 163]. Typically, these transporters consist of a single membrane domain that catalyzes the transport activity, which can be bidirectional (Fig. 1a) [129, 163]. The major facilitator superfamily (MFS) contains the best studied examples of secondary transporters known to date, and is therefore taken as a general model. Intriguingly, the structure of these transporters is remarkably uniform, despite the lack of amino acid sequence homology.

### *2.1 Modes of transport*

Transport mediated by secondary transporters can be divided into three major categories (Fig. 2) [163]. First, when uptake of a solute is independent of a coupling ion, the process is termed 'uniport' (Fig. 2a). Solute uniport can be electrogenic when the solute 'S' is cationic, in which case uptake is driven both by the chemical gradient of the solute ($\Delta\mu_S$) and the transmembrane potential ($\Delta\Psi$), as is the case for L-lysine and L-arginine uptake by *Bacillus stearothermophilus* [75]. Alternatively, when the solute is uncharged, uniport is electroneutral and the driving force is provided by the $\Delta\mu_S$ alone. During 'symport', the second transport mode, the uptake of a solute 'S' requires co-transport of one or more ions, such as $H^+$ or $Na^+$ (Fig. 2b). Symport can either be electrogenic or electroneutral, depending both on the charge of the transported solute and the charge and number of co-transported ions. These charactistics in turn determine the relative contributions of the proton gradient ($\Delta$pH), as well as $\Delta\Psi$ and $\Delta\mu_S$ to the driving force. Well-studied examples of prokaryotic symporters are the

galactoside-$H^+$ symporter LacY; the rhamnose-$H^+$ symporter RhaT; and the fucose-$H^+$ symporter FucP; all of *E. coli* [94, 145, 196].

Finally, when the transporter translocates substrates in opposing directions across the membrane, the process is termed 'antiport' or 'countertransport' (Fig. 2c). In the case of substrate-ion antiport the driving forces can vary drastically with the charge of the transported solute, as the overall transport process is either electrogenic or electroneutral (reviewed extensively in ref: 163). In a number of cases, the antiporter facilitates precursor/product exchange in a metabolic reaction. Here, both the inward gradient of the substrate 'S' and the outward gradient of the product 'P', which are essentially maintained by the metabolism, work together to drive transport. Depending on the charges of the respective transported compounds, the transport process in some cases generates rather than consumes the pmf [163]. This is the case for citrate uptake during citrolactic fermentation by the lactic acid bacterium *Leuconostoc mesenteroides* [136].

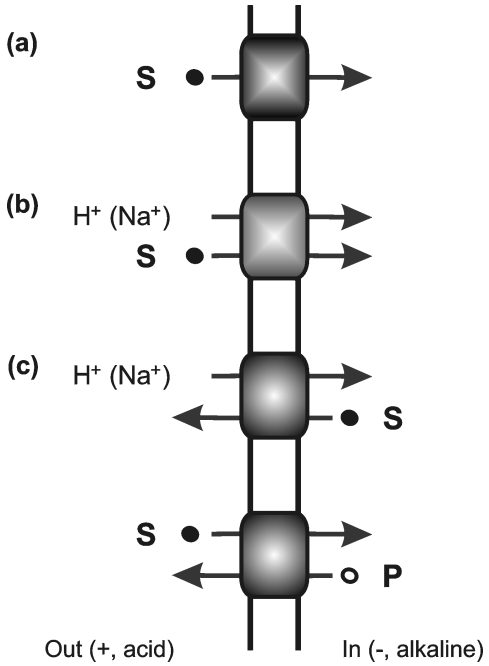



**Figure 2.** Three modes of transport catalyzed by secondary transporters (see text for details). (a) Uniport. Transport of the solute is independent of a coupling ion. (b) Symport. Transport of the solute requires co-transport of one or more ions, such as $H^+$ or $Na^+$. (c) Antiport. The transporter translocates solutes in opposing directions across the membrane. S: substrate; P: product

*2.2 Model secondary transporters: the Major Facilitator Superfamily*

The most prominent family in the class of secondary transport proteins is the major facilitator superfamily (MFS) [65, 132, 157]. This family comprises over 1000 fully sequenced members divided over at least 17 subfamilies [157] (for an updated overview, see: http://www-biology.ucsd.edu/~msaier/transport/). MFS transporters are involved in the uptake and efflux of a broad range of substrates [157]. Despite this diversity members can be identified via the conserved signature sequence G-[RKPATY]-L-[GAS]-[DN]-[RK]-[FY]-G-R-[RK]-[RKP]-[LIVGST]-[LIM] (all possibilities for a single position are given between brackets) in the cytoplasmic loop connecting transmembrane domains (TMDs) 2 and 3 [157, 163]. This stretch of amino acids has been suggested to be involved in global conformational changes occuring during the transport event, and some of the residues have been shown to be essential for transport activity. Interestingly, a similar motif is found in the loop connecting TMDs 8 and 9, but these residues are not essential [157, 163]. This finding has led to the suggestion that these transporters have evolved from a common 6-TMD ancestor by a gene-duplication event [180].

*2.3 Structure: the 12-transmembrane domain model*

Amino acid sequence analysis of the majority of described MFS family members, using hydropathy profiling as well as the 'positive inside' rule postulated by Von Heijne, has resulted in a general structural model [74, 226, 227]. According to this model, most members of the MFS consist of a single membrane domain with 12 $\alpha$-helical TMDs traversing the membrane in a zigzag manner. In this topology, both the N- and the C-terminus are located in the cytoplasm, and a large, central cytoplasmic loop connects TMD 6 and TMD 7 [65, 157, 163]. Some MFS members have been described that span the membrane 14 or 24 times, but these can be truncated into functional 12-TMD units [92, 240]. Occasionally, odd numbers of helices like 11 or 13 TMDs are found. Taken together, these findings indicate the 12-TMD topology represents the functional transporter unit.

Several models exist that address the packing of these helices within the membrane. A general model was proposed based on computational sequence analysis, where the length of the cytoplasmic loop between TMDs 6 and 7 allows packing of both helices at opposing positions outside of the helices that are lining a single central pore [65]. A second model was based on the immense body of structural data available on the lactose permease LacY of *E. coli* [194]. The charge distribution along the respective TMDs, interaction sites between helices as mapped by cysteine scanning mutagenesis and predicted length of both the periplasmic and cytoplasmic loops have refined the MFS model for helix packing. This LacY-model shows two cavities that are accessible from the extracellular side, and one cavity that is accessible from the cytoplasmic side [194].

For a few secondary transporters two-dimensional crystals have been obtained that at their limited resolution provide a first glimpse of helix packing. These are the $Na^+/H^+$ antiporter NhaA of *E. coli* [235, 236], the melibiose permease MelB of *E. coli* [69], and the oxalate/formate antiporter OxlT of *Oxalobacter formigenes* [81]. The transmembrane helices of the $Na^+/H^+$ antiporter NhaA are organised in a two-times-six grouping, but both sub-domains are asymmetrical with regard to helix packing [235, 236]. Whether the protein is present as a functional dimer or a monomer cannot be concluded from the available data [235]. The non-related melibiose transporter MelB also crystallizes in an asymmetrical protein unit, with two sub-domains lining a curve-shaped cleft, and also for this protein the functional mono- or multi-meric state could not be inferred [69]. It has been suggested that this asymmetry is functional in $Na^+$-coupled transporters, but such statements await further investigation. In contrast, the oxalate/formate antiporter OxlT crystallyzes with remarkable intramolecular symmetry, with two subunits of 6 TMDs surrounding a large central cavity [80, 81]. This central cavity is widest at the center of the membrane, and it has been suggested that this structure resembles the intermediate state in substrate transport, according to a mechanism where one substrate-binding site alternatively is exposed to the cytoplasm or the extracellular milieu [81]. The protein is presented as a functional monomer. The reconstituted three-dimensional structure was used to propose a general model for MFS transporter proteins, which differs from the one proposed for LacY [80]. In the OxlT model, the 12 TMDs are arranged symmetrically around a central cavity, with TMD 6

and TMD 7 at opposing sides. The substrate is proposed to pass through this central cavity. Several (parts) of the helices line the central cavity with a large variety in contributing residues, which thereby allows for the broad range of possible transported substrates within the family [80]. Finally, the striking two-fold symmetry of the protein, with 6 helices in each half of the monomer, seems to support the view that MFS transporters may have evolved from a 6-TMD ancestor by a gene duplication event [180].

Although the structural data obtained with these proteins thus far is in accordance with the 12-TMD model for MFS transporters, differences in helix packing and intramolecular symmetry are observed. In addition, it should be noted that not all secondary solute uptake systems with predicted 12-TMD topology actually have such a topology. The sodium ion-dependent citrate carrier CitS of *Klebsiella pneumoniae* was shown to traverse the membrane 11 times, instead of 12 [216]. Apart from improvements in structural resolution which would facilitate even more accurate helix assignment, the importance for the combinatorial approach used for LacY is evident when attempts are made to extrapolate findings to other putative 12-TMD secondary transporters. Although CitS is not a member of the MFS, but of the 2-hydroxycarboxylate transporter family [28], which is referred to as the citrate:cation symporter family (2.A.24) at http://www-biology.ucsd.edu/~msaier/transport/, the results obtained with this protein clearly show that statements concerning membrane protein topology should preferably be based on experimental as well as structural data.

## 3. Binding protein-dependent secondary transporters

The second major class of transport proteins is that of the binding protein-dependent secondary transporters [53, 59, 98, 171] (Fig. 1b). Members of this family so far have only been found in prokaryotes. This class of transporters has been labeled tri-partite ATP-independent periplasmic (TRAP) transporters by Kelly *et al.* [59, 98, 171]. However, in a number of family members both membrane domains have 'fused' into a single polypeptide (Fig. 1b), and we therefore prefer the name 'binding protein-dependent secondary transporters' in analogy with the terminology used to identify the other transport systems.

*3.1 Family portrait*

Transport by members of this family is mediated by a high-affinity substrate-binding protein that delivers the solute to an integral membrane permease (Fig. 1b). Examples of the substrate-binding protein are the DctP protein of *Rhodobacter capsulatus* [59], and YiaO of *E. coli* [Chapter 4]. Amino acid sequence alignment of the known substrate-binding proteins revealed a potential signature sequence: [SANF]-[RKNSQ]-x(2)-[LIV]-x(2)-[LIVWPF]-[ADES]-[DHE]-[LIVMF]-x-G-[LGM]-[KRD]-[LIVM]-R-[LIVTA] (residues between brackets: conserved substitutions; 'x(n)': number of 'n' random amino acid substitutions) [171]. This sequence retrieved only members of this family when screened against the SWISSPROT and TREMBL databases, and none of the solute-binding proteins from members of the ABC superfamily [171].

The permease domain consists of two structural subunits. The large subunit, for example DctM of *Rh. capsulatus* [59] or YiaN of *E. coli* [Chapter 4], resembles the secondary MFS(-like) proteins. Amino acid sequence alignment identified what may be seen as a signature sequence for these proteins in the TMD 2-region: [LIVM]-[LIVMT]-A-[LIV]-P-]FL]-[FY]-[LIVMA]-x(2)-[GS]-x(2)-[LM]-x(2)-[GST]-x(2)-[GAST]-x(2)-[LIVM] [171]. The second, smaller subunit consists of 4 putative TMDs and shows no homology to known membrane proteins [53, 98, 171]. Examples of these proteins are the DctQ protein of *Rh. capsulatus* [59] and the YiaM protein of *E. coli* [Chapter 4]. A signature sequence has also been identified for these proteins: [LIVC]-F-x-[WLY]-[LIVFG]-G-[AI]-[LIVAS]-[LIVY]-x(6)-H-[LIVM]-x-[LIV], in the TMD 2 region of the protein [171]. Mechanistic details of binding protein-dependent secondary transport are lacking at this time, therefore the function of these conserved amino acid stretches is unknown. The ABC-transporter signature sequence (see section 4) is not found in either membrane domain [53, 59, 171].

*3.2 Physiological function*

Interestingly, these transporters are found throughout both prokaryotic Kingdoms of life, i.e. the Bacteria and the Archaea, indicating that it can be considered an ancient transport protein family [54, 171]. It is therefore surprising that the number of family members found per completely sequenced genome is relatively low, i.e. ranging from one in the *E. coli* K-12 genome [22, Chapter 4] up to 12 in the recently sequenced *Bordetella* genome [10]. In comparison, the *E. coli* K-12 genome contains 44 binding protein-dependent ABC transporters [117] and up to 71 members of the MFS (listed at http://www-biology.ucsd.edu/~paulsen/transport/; 'Transport proteins in *Escherichia coli*'). In light of this, functional analysis of each member is extremely important, as it will increase our understanding of the entire transport protein family.

The first member of this transport protein family was characterized at a biochemical level in the purple photosynthetic Gram-negative bacterium *Rhodobacter sphaeroides*. The transporter was induced by sodium in the medium, and functions as a binding protein- and pmf-dependent uptake system for the anionic amino acid glutamate (Fig. 3b) [91]. The first binding protein-dependent secondary transporter to be characterized both at the functional and the molecular level, DctPQM of *Rh. capsulatus*, was shown to transport the $C_4$-dicarboxylates malate, succinate and fumarate (Fig. 3a) [59, 73, 190]. Homologs have been found that transport $C_4$-dicarboxylates in *Wolinella succinogenes* [210] and sense $C_4$-dicarboxylates in *Bacillus subtilis* [13]. In addition, a glutamate transporter was identified in *Synechocystis* sp. [169], and based on genomic organisation, the FN1472 and FN1473 proteins of *Fusobacterium nucleatum* have been tentatively designated as a *N*-acetylneuraminate (Fig. 3d) uptake system [95]. This suggested that all family members function as organic anion transporters. However, the increase in experimental evidence is now starting to show the diversity in possible functions of this class of transporters. The osmoregulated TeaABC transporter of the halophilic bacterium *Halomonas elongata* was shown to transport the zwitterionic compatible solutes ectoine and hydroxyectoine (Fig. 3c) [66]. This finding expands the possible function of this type of transporter from carbon source uptake to protection of the cell against unfavourable extracellular conditions.

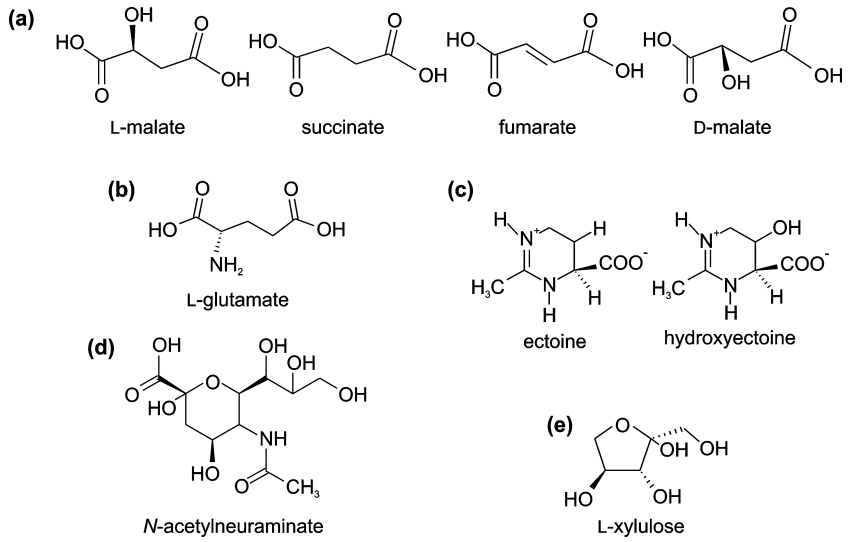



**Figure 3.** Nine compounds that thus far have been identified as substrates of binding protein-dependent secondary transporters. (a) DctP of *Rh. capsulatus* binds L-malate > succinate > fumarate > D-malate [232]. (b) An as of yet unidentified glutamate-binding protein, perhaps encoded by Rsph_p_327 (see section 3.3), was found in *Rh. sphaeroides* [91]. (c) TeaA of *H. elongata* binds the compatible solutes ectoine and hydroxyectoine [66, 206]. (d) Based on genomic organization, the FN1472 protein of *F. nucleatum* is thought to bind *N*-acetylneuraminate [95]. (e) The YiaO protein of *E. coli* K-12 has recently been shown to bind L-xylulose [Chapter 4].

More recently, we have shown that the YiaMNO transporter of *E. coli* K-12 is involved in transport of the rare pentose L-xylulose (L-*threo*-2-pentulose, Fig. 3e) [Chapter 4]. The physiological function of this uptake activity is as of yet unknown [Chapter 3], but our findings confirm that members of this transport protein family are capable of transporting substrates other than organic anions (Fig. 3).

### *3.3 The substrate-binding domain*

Thus far, none of the substrate-binding domains from the family of binding protein-dependent secondary transporters has been crystallized. Since these proteins are not homologous to solute-binding domains of the ABC superfamily (see below and ref: 171), it will be of interest to determine if these proteins share structural similarity and how they are evolutionary related (see section 4.2). These proteins are synthesized as

precursors with an N-terminal signal sequence, and it is generally assumed that they are exported to the periplasm of Gram-negatives via the Sec-pathway [160, 206].

A basic local alignment search tool (BLAST) search of the databases available at http://www.ncbi.nlm.nih.gov/BLAST/, using the amino acid sequence of the L-xylulose-binding protein YiaO of *E. coli* K-12, revealed the presence of 50 homologous sequences in over 20 prokaryotes. None of the extramembranal substrate-binding domains of the ABC transporter superfamily were identified by this BLAST search. The amino acid sequences of 28 close homologs (listed in Table 1) were aligned using the ClustalW program available at http://www.ebi.ac.uk/clustalw. The sequence comparison was performed with default opening, end, extending and separation gap penalties (10, 10, 0.05 and 8, respectively). The phylogram data were applied in the Phylip Drawtree program (Phylogeny Interference Package version 3.5c, University of Washington) and used to construct an unrooted tree diagram. The phylogenetic tree obtained in this manner clearly shows that these proteins cluster in what may represent functionally homologous groups (Fig. 4). The proteins involved in $C_4$-dicarboxylate-binding and -sensing (labelled Rcap, Wsuc, and Bsub, see Table 1) are found in one domain of the tree, whereas the YiaO protein (labelled Ecol) clusters in a clearly distinct domain with three close homologs that may specify similar functions (Fig. 4). The ectoine/hydroxyectoine-binding protein of *H. elongata* (labelled Helo) localizes in the third major domain of the tree, as does the putative *N*-acetylneuraminate transporter of *F. nucleatum* (labelled Fnuc). Interestingly, the two homologs that were detected in *Rh. sphaeroides*, labelled Rsp1 and Rsp2 in Fig. 4, localize in the domain that contains the $C_4$-dicarboxylate-binding and -sensing proteins. These proteins likely represent the glutamate- and succinate/malate-binding proteins of this organism [91]. The DctP protein of *Rh. sphaeroides* (Rsp1, Fig. 4) is a close homolog of DctP of *Rh. capsulatus* and is likely a succinate/malate-binding protein. This leaves Rsp2 as a possible candidate for the glutamate-binding activity described before [91]. The GtrABC transporter of *Synechocystis* sp. strain PCC6803 functions as a sodium-dependent glutamate transporter [169] but does not cluster with the aforementioned organic anion-binding proteins (not shown) [160].

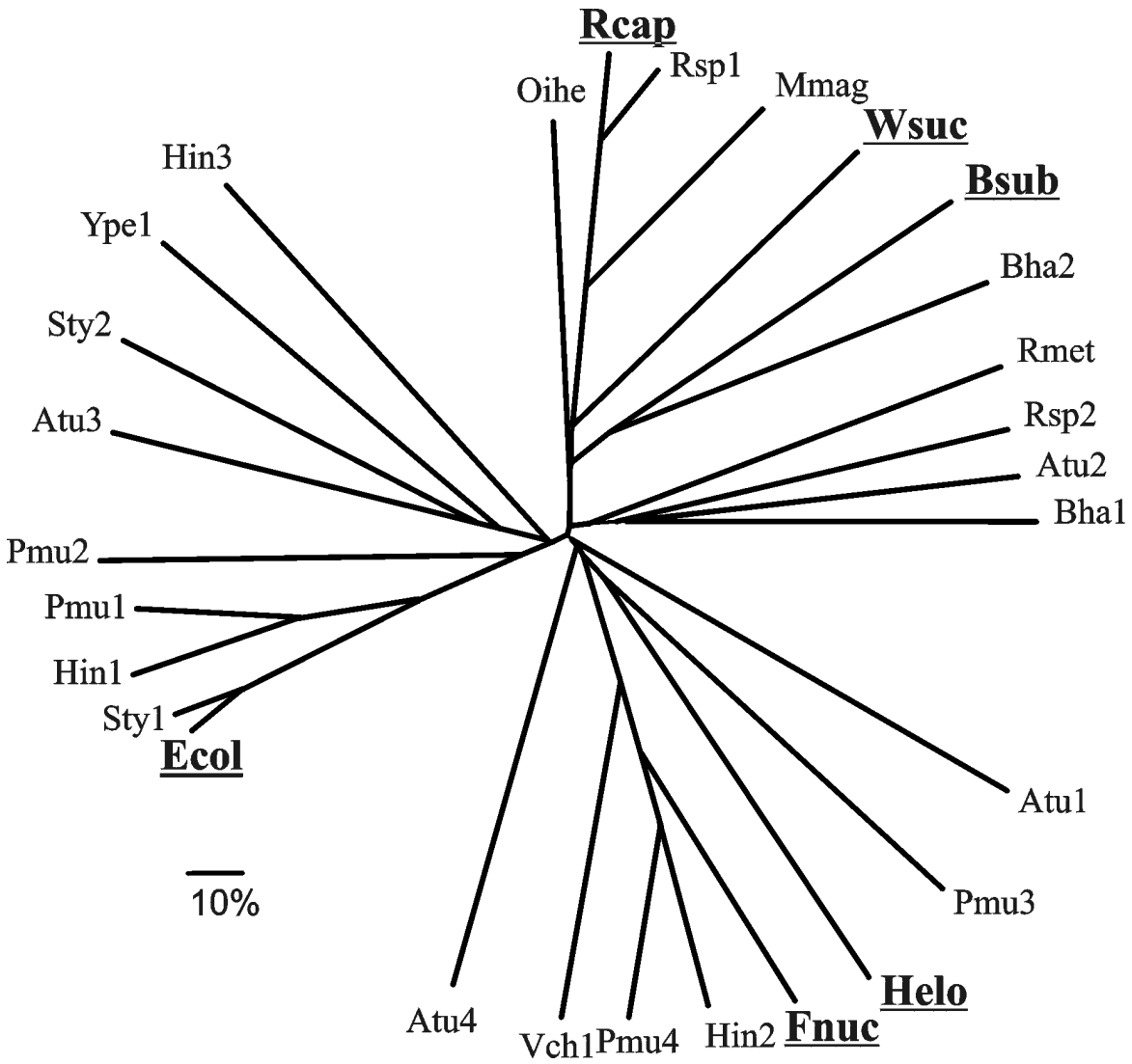



**Figure 4.** Amino acid sequence alignment of 28 homologs of the *E. coli* YiaO protein (listed in Table 1) reveals what may be functional groups. The amino acid sequences were compared using the ClustalW program with default opening, end, extending and separation gap penalties (10, 10, 0.05 and 8, respectively). The tree was created using the Phylip Drawtree program, where the distances represent the differences in identical amino acid residues in percentages.

**Table 1.** Selected YiaO-homologs used to construct Fig. 4

| Name: | Organism: | Locus: | Size[a]: | E-value[b]: | Label[c]: |
|---|---|---|---|---|---|
| YiaO | *Escherichia coli* K-12 | AAB18556 | 328 | | Ecol |
| YiaO | *Salmonella typhimurium* LT2 | AAL22532 | 328 | -165 | Sty1 |
| PM1252 | *Pasteurella multocida* | AAK03336 | 328 | -108 | Pmu1 |
| HI1028 | *Haemophilus influenzae* Rd | AAC22688 | 328 | -102 | Hin1 |
| PM1644 | *P. multocida* | AAK03728 | 325 | -56 | Pmu2 |
| DctP | *Agrobacterium tumefaciens* C58[d] | AAK90228 | 341 | -48 | Atu1 |
| AGR_L_3344 | *A. tumefaciens* C58 | AAK90249 | 282 | -47 | Atu3 |
| BH2673 | *Bacillus halodurans* | BAB06392 | 348 | -46 | Bha1 |
| OB0243 | *Oceanobacillus iheyensis* | BAC12199 | 337 | -45 | Oihe |
| DctP | *Rhodobacter sphaeroides* | AF005842 | 335 | -44 | Rsp1 |
| AGR_C_4976 | *A. tumefaciens* C58 | AAK88459 | 335 | -43 | Atu2 |
| PM1651 | *P. multocida* | AAK03735 | 334 | -41 | Pmu3 |
| FN1472 | *Fusobacterium nucleatum* ATCC 25586 | AAL95665 | 327 | -40 | Fnuc |
| DctP | *Wolinella succinogenes* | CAA10756 | 329 | -39 | Wsuc |
| Rsph_p_327 | *Rh. sphaeroides* | ZP_00004447 | 331 | -38 | Rsp2 |
| HI0146 | *H. influenzae* Rd | AAC21818 | 329 | -37 | Hin2 |

**Table 1.** Selected YiaO-homologs used to construct Fig. 4 (*continued*)

| Name: | Organism: | Locus: | Size[a]: | E-value[b]: | Label[c]: |
|---|---|---|---|---|---|
| VC1779 | *Vibrio cholerae* N16961 El Tor | AAF94928 | 321 | -35 | Vch1 |
| YPO1579 | *Yersinia pestis* | CAC90401 | 330 | -35 | Ype1 |
| Reut_p_362 | *Ralstonia metallodurans* | ZP_00021455 | 354 | -34 | Rmet |
| YdbE | *Bacillus subtilis* | BAA19281 | 350 | -34 | Bsub |
| HI0052 | *H. influenzae* Rd | AAC21730 | 328 | -33 | Hin3 |
| PM1709 | *P. multocida* | AAK03793 | 327 | -33 | Pmu4 |
| BH2750 | *B. halodurans* | BAB06469 | 348 | -33 | Bha2 |
| STM3169 | *S. typhimurium* LT2 | AAL22043 | 327 | -32 | Sty2 |
| DctP | *Rhodobacter capsulatus* | CAA45385 | 333 | -32 | Rcap |
| AGR_L_115 | *A. tumefaciens* C58 | AAK88621 | 347 | -32 | Atu4 |
| Magn_p_8881 | *Magnetospirillum magnetotacticum* | ZP_00056052 | 352 | -32 | Mmag |
| TeaC | *Halomonas elongata* | Unknown [e] | 340 | n.d.[f] | Helo |

[a] number of amino acids

[b] protein BLAST using *E. coli* K-12 (MG1655) YiaO as query

[c] identifier in Fig. 4

[d] sequence from Cereon, identical to that from U.Washington

[e] sequence provided in ref. 206

[f] not determined

*3.4 The permease*

The permease of a binding protein-dependent secondary transporter consists of two structural domains (Fig. 1b). The hydropathy profile of the large subunit is remarkably similar to MFS protein structures. 12 putative TMDs traverse the membrane in a zigzag manner with both termini located in the cytoplasm. A large cytoplasmic loop separates TMD 6 from TMD 7 [53, 98, 171]. This resemblance to 'classical' secondary transporters has led to the suggestion that this domain is involved in energy coupling and transport. The second subunit consists of 4 putative TMDs and is not homologous to known membrane proteins [53, 98, 171]. Topological analysis has confirmed the number of TMDs and the position of both termini in the cytoplasm in the case of the DctQ protein of *Rh. capsulatus* [241]. This domain may be involved in the interaction with the substrate-binding protein or coupling to the putative transporter domain (see section 5.4).

Interestingly, whereas the 'classical' secondary transporters are able to mediate bi-directional transport [163], activity of at least the glutamate transporter of *Rh. sphaeroides* appears strictly unidirectional and cannot be detected in the absence of the solute-binding domain [91]. It is not certain if this is a general property of these transporters as this has not been studied in detail for the other reported systems. It would be interesting to investigate whether the large, MFS-like membrane domain by itself is able to function as a secondary transporter. If this were the case, the observed directionality of the multicomponent system may be imposed by either the substrate-binding domain or the small membrane domain, or both.

A recent genomic databasis analysis demonstrated the existence of ABC transporters with the extramembranal substrate-binding domain fused to the integral membrane component [213]. Strikingly, the amino acid sequence of one putative binding protein-dependent secondary transporter, Y4mM of *Rhizobium* sp. strain NGR234 [61] (GenBank accession number 16519640) shows a similar structural organization. The N-terminal part of the protein is homologous to substrate-binding proteins of this transporter family, whereas the C-terminus is homologous to the small, 4 TMD membrane subunit (Fig. 5a) [61].

However, no signal sequence for secretion into the periplasm was detected using the SignalP program [149], and the putative substrate-binding domain is predicted to be in the cytoplasm (Fig. 5b) with the hydropathy profiling program available at http://www.cbs.dtu.dk/services/TMHMM-2.0/ (data not shown). Therefore, it is unlikely that this is a functional configuration.

**(a)**

1 MKYEITRR**RFLAASSAVLAAPAIVTMVRPARAGTTLTLGH**
41 **GAAPGNPRTVAAAKFAELVAEKTAGRVTINVAGAETLGSD**
81 **AAMLTSLRTGALDVTANSQGATSALVPELAALGLPFLFEN**
121 **TAKAMQVLGGPVGAELVKRFEAVGVVPLDWWDNGIRHLTN**
161 **SKRKVAAPAEVSGMKIRTPADPMTMDIFQALGAATEQIAF**
201 **GELYVALQQGVVDGQENPLANIDSSKLYEVNKYISLTGHK**
241 **WESTPFLMSQIAQARLGGDLEAVKAAAKEAGELQRKLSAD**
281 **KDAEVLAAFRRISAIEVTEVDREGFAKATASVVESRRSPS**
321 **GISSPRSNRQPKAEALSAREHPM**KSLSNLVE*LTARAIVWF*
361 *ARQVVIFSGIALMVFMTANVAARYVLAGGGFSFAQELPVL*
401 *IFPWFILGGIVLAAHSGGHMAVEWIYDKLRDGARSTAFVA*
441 *ANLVSAGAFLMLGYQAYLVGEIAGIEHSPVLQLPNSVGYF*
481 *ALAVGSVLVAIVTLAVALRVLRLGWDHRTNTESG*EVAL

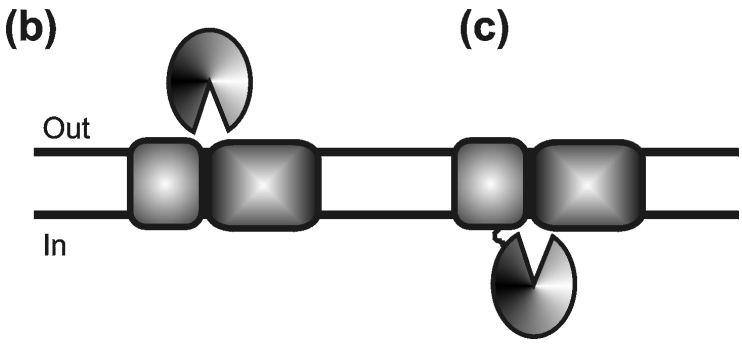



**Figure 5.** (a) Amino acid sequence of Y4mM of *Rhizobium* sp. strain NGR234 [61] (GenBank accession number 16519640). The N-terminus (bold) is homologous to the substrate-binding domains, and the C-terminus (italics) to the small membrane domain of the binding protein-dependent secondary transporters, respectively. (b) Structural model of a binding protein-dependent secondary transporter: the substrate-binding protein is located on the outside of the cell. (c) Structural model of Y4mM: the absence of a secretion signal sequence, as well as the TMHMM program (see text) predict the N-terminus to be in the cytoplasm which most likely is a non-functional structure.

## 4. Binding protein-dependent primary transporters

Binding protein-dependent primary transporters (Fig. 1c) are multicomponent, high-affinity, ATP-driven uptake systems that are all members of the ABC transporter superfamily. This family is found widespread throughout all kingdoms of life, and several members have in humans been linked to disease [7, 26, 40, 57, 77, 79, 86].

### *4.1 The ABC superfamily portrait*

Some members of the ABC transporter superfamily are uptake systems (importers) whereas others are excretion systems (exporters). Combined they transport an enormous range of potential substrates with a large diversity in molecular characteristics [40, 57, 77]. Despite this diversity in physiological function, the overall domain composition is remarkably similar [26, 78, 86]. These transporters consist of two homologous or identical permease domains, and two homologous or identical NBDs (Fig. 1c). These structural domains are expressed either as separate polypeptides or fused in various combinations [7, 40, 77, 78, 86]. All the structural components of binding protein-dependent ABC transporters are required for transport activity [47, 111, 118].

Members of the ABC superfamily are identified through a highly conserved region of the NBDs that has been labeled the 'ATP-binding cassette': the Walker A (G-x(2)-G-x-G-K-S-T) and Walker B (h(4)-D-E-P-T; 'h': hydrophobic residue) motifs [230], separated by a stretch of amino acids containing the ABC signature sequence L-S-G-G-Q [77, 86, 117]. These sequences are directly involved in nucleotide binding (see section 4.4). The membrane domains of the binding protein-dependent ABC transporters can be identified via the conserved E-A-A-A-x(3)-G-x(9)-I-x-L-P motif, also known as the 'EAA-region', at approximately 100 residues from the C-terminus [41, 185]. This conserved EAA-region is localized in a cytosolic loop [186] and is involved in the interaction with the NBDs (see section 5.2). Interestingly, detailed sequence analysis has shown that substrate-binding proteins that are specific for the same or very similar compounds show a higher degree of sequence similarity than substrate-binding proteins specific for very different compounds but originating from the same organism [200].

*4.2 The substrate-binding domain*

Due to the relatively easy accessibility of the extramembranal substrate-binding proteins of ABC transporters both in bacteria and archaea, these proteins have been a major topic in research over the past 30 years. Starting with the L-arabinose-binding protein of *E. coli* [134] an increasing number of these proteins has been crystallized, and the respective three-dimensional structures have been elucidated. Despite the lack of overall amino acid sequence homology, the substantial size difference (ranging from 20 to 60 kDa), and the wide range of potential substrates, the overall structures of these proteins are very similar [7, 62, 170, 200]. In general, the substrate-binding proteins consist of two distinct globular domains that are built up mainly of alternating $\beta$-sheets and $\alpha$-helices, and are connected by a more [27] or less [62, 84, 170] rigid 'hinge' region. Differences in the number of connecting strands in this hinge region have in the past led to the classification of substrate-binding proteins in two sub-groups. In group I proteins, e.g. MBP of *E. coli*, three short segments connect both globular domains, whereas in the members of group II, e.g. the oligopeptide-binding protein $OppA^{Sty}$ of *S. typhimurium*, two connecting segments are found [170]. More recently, structures were determined that showed the presence of a single, rigid $\alpha$-helix in this hinge region of the vitamin $B_{12}$-binding BtuF protein of *E. coli* [27, 97], which may therefore specify a different sub-group according to this classification.

In all structures determined thus far, the substrate is bound in the deep cleft separating both lobes that engulf the molecule during binding by a hinge-bending motion of the connecting region (see section 5.1) [72, 128]. This movement of domains is less pronounced in binding proteins with a more rigid hinge region such as BtuF [97].

*4.3 The permease*

The integral membrane subunits of prokaryotic binding protein-dependent ABC transporters are not homologous to those of prokaryotic exporters or eukaryotic ABC transporters [185]. Sequence analysis of these permeases has revealed that they cluster in groups that in general transport related compounds [185, 186], in analogy to the substrate-binding domains of these transporters [200]. Similar analyses using the amino

acid sequences of the ABC domains provide a more detailed subdivision of these clusters [40, 110].

The canonical ABC transporter permease consists of two homologous or identical membrane domains, each built up of 6 α-helical TMDs [57, 77, 86]. In recent years, various attempts have been made to determine the three-dimensional structures of these permeases in order to test this hypothesis. LmrA of *L. lactis* is an ABC transporter responsible for multidrug resistance in this Gram-positive bacterium [218]. The *lmrA* gene encodes a 'half transporter' of one membrane domain, of 6 putative transmembrane α-helices, and one cytosolic NBD [218]. Polarized attenuated total reflection infrared spectroscopy confirmed the α-helical nature of the transmembrane domains [67]. Cysteine scanning mutagenesis and subsequent accessibility studies supported the presence of 6 TMDs in each monomer of LmrA. In addition, these studies detected an aqueous chamber within the membrane domain under non-energized conditions that is open to the cytosol [161].

Low-resolution crystal structures of the multidrug-resistance protein MDR1, also referred to as P-glycoprotein or P-gp (at 2.5 nm resolution) [178], and of the transporter associated with antigen processing TAP [221] have yielded information on the subunit composition of these eukaryotic ABC transporters. MDR1 was purified as an active monomer, and although no helices could be assigned the data were in support of the 12 TMD hypothesis for the complete transporter. In addition, a large, possibly aqueous pore was suggested to exist within the membrane domain [178]. TAP was isolated as an active heterodimer of the two 'half transporters' TAP1 and TAP2 [221]. At low resolution this ABC transporter shares three-dimensional features with P-gp, i.e. it consists of two membrane domains and two NBDs [221]. Nevertheless, biochemical studies have in the past revealed the presence of eight (TAP1) and seven (TAP2) transmembrane α-helices in the respective subunits [228], and others have proposed a number of ten (TAP1) and nine (TAP2) [113]. Some of these 'additional' helices may be required as targeting signals for correct localization of the protein in the ER [113].

Only recently, high-resolution crystal structures of complete ABC transporters have become available. The first to be published was that of the lipid A transporter MsbA of *E. coli*, at 4.5 Å resolution [34]. Although the structure was obtained in the absence of a lipid membrane, it shows the functional transporter as a homodimer of the half

transporter encoded by the *msbA* gene. The membrane domain of each monomer consists of 6 α-helices that all are tilted between 30º and 40º from the plane of the membrane. As the monomers contact each other only in the outer membrane leaflet, a cone-shaped structure is formed with substantial openings to either side facing the lipid bilayer. In addition, a cone-shaped chamber is formed within the membrane domain, which may be lipid-filled *in vivo* [34]. The presence of a large chamber within the membrane, which may be either lipid-filled or aqueous, has been reported in the protein structures of the (functional) MsbA homologs MDR1 [178] and LmrA [161]. Cryo-electron spectroscopy has shown that the half transporter YvcC of *B. subtilis* is structurally very similar to MsbA [33]. As these proteins are all thought to transport rather hydrophobic compounds from the inner leaflet of the lipid bilayer of the membrane [25], this structural organisation probably reflects the specific requirements for this activity. The NBDs of each monomer are located at the cytoplasmic side of the membrane, and show significant structural homology to HisP, the NBD of the His transporter of *S. typhimurium* (see section 4.4) [34, 85]. Remarkably, the transmembrane domains are connected to the nucleotide-binding domains by two intracellular domains (ICDs) which may be distinctive features for members of the MDR-subfamily of ABC transporters (see section 5.2) [34]. However, uncertainty exists concerning this configuration of the subunits, as biochemical evidence strongly suggests that the NBDs are in close proximity. Possibly, MsbA has been crystallized as a monomer, while the crystallographic dimer may not be of physiological significance.

The first crystal structure of a binding protein-dependent ABC transporter was published more recently, and shows a rather different organisation of the respective domains. The structure of the vitamin $B_{12}$ transporter BtuCD of *E. coli* was resolved at 3.2 Å resolution [122], in the absence of its cognate substrate-binding domain BtuF [30]. The functional unit indeed consists of two copies of both the membrane domain BtuC and the NBD BtuD. Both NBDs are in close contact, as are the transmembrane domains, which is in strong disagreement with the MsbA structure described above (see section 4.4) [122]. Another remarkable difference is the number of transmembrane α-helices, which in this structure amounts to 10 in each BtuC subunit, and 20 in the complete transporter [122]. This number compares to what has been proposed for TAP [113]. Since the membrane component of ABC transporters is less conserved than the

ABC signature in the NBDs, the observed differences between members of the ABC superfamily probably relate to differences in transport function.
The helices of the BtuC subunit are packed in an intricate fashion, and the membrane domain itself is positioned perpendicular to the plane of the membrane, as opposed to the MsbA structure [122]. The long periplasmic loop between transmembrane helices 5 and 6 may be involved in binding of BtuF. Together, the membrane domains of the homodimer probably line the translocation path, leaving a cavity to the periplasmic face of the membrane that would suffice to accommodate a vitamine $B_{12}$ molecule. The translocation path appears to be 'gated' at the cytoplasmic side [122].

Another important structural feature was discovered in the cytoplasmic loop in between transmembrane helices 6 and 7, which consists of two short helices connected by a sharp bend, resembling an 'L' shape. This 'L-loop' contains the EAA-region of bacterial binding protein-dependent ABC transporters. A similar intracellular loop of the cystic fibrosis transmembrane conductance regulator CFTR has been shown to be functionally important [122]. The L-loop may represent a general interface between the membrane domain and the NBD of the ABC transporter (see section 5).

Taken together, it appears that the structure of the permease may vary not only with the type of substrate it transports, but also with the direction in which it transports its substrates. This conclusion may however be too premature as the limited availability of high-resolution crystal structures at this time is ground for caution. The binding protein-dependent ABC transporters most likely belong to a different subclass of this superfamily than the ABC type exporters (see 'ABC transporters' classified under 3.A.1 at http://www-biology.ucsd.edu/~msaier/transport/). Overall, the transport protein crystals have provided a wealth of information on the structures and thereby possible transport mechanisms of these proteins (see section 5). It is interesting, however, to note that in order to obtain the two high-resolution structures that are available to date, over 50 full-length ABC transporters have been cloned and overexpressed [34, 122], underlining the difficulty of the method.

*4.4 The nucleotide-binding domain*

The NBDs of ABC transporters are the only structural components that show significant overall sequence homology within the ABC transporter superfamily [40, 79, 86, 188]. The family derives its status from this homology which identifies a highly conserved region of the NBDs that has been labeled the 'ATP-binding cassette', containing the Walker A and B motifs, as well as the ABC signature sequence L-S-G-G-Q [77, 86, 117, 188, 230]. Additional stretches of conserved amino acid residues have been identified, i.e. the 'P-loop' that encompasses the Walker A motif [184], the 'switch region', and the 'Q-loop' [83]. The function of each of these regions in ATP-binding and -hydrolysis, as well as signalling to the transmembrane domain, is under continuous investigation. The involvement of the NBD in ATP binding and hydrolysis has been well documented [4, 44, 150, 222, 242].

In recent years, high-resolution crystal structures have been obtained for the NBDs of several of the ABC transporters. The first report was on HisP, the NBD of the histidine transporter of *S. typhimurium* [85]. However, its validity as a general model for the NBDs of ABC transporters is being challenged by the more recent structures that have become available. The ABC-fold is similar to that of MalK of *Thermococcus litoralis* [50], the NBD of MsbA [34], and that of the DNA double strand repair protein Rad50 of *Pyrococcus furiosus*, which is not part of a permease but does contain an ATP-binding cassette [83]. However, the dimeric organization of HisP in the crystal has been questioned [50]. The conserved sequences that are involved in ATP-binding and -hydrolysis by the NBD are present in three subdomains in the *Methanococcus jannaschii* MJ1267 protein: i) the $F_1$-type ATP-binding core containing the Walker A and B motifs, ii) the antiparallel $\beta$ sheet subdomain involved in positioning of the base and ribose moiety of the nucleotide, and iii) the $\alpha$-helical subdomain containing the ABC signature [96]. Another striking feature is found in MalK of both *T. litoralis* and *E. coli*, which both contain an additional domain that has been linked to a function in regulation [50], and a similar domain has been found in GlcV of *S. solfataricus* [222].

The NBDs of ABC transporters are believed to be functional as dimers. MalK, the NBD of the maltose transporter of *E. coli*, dimerizes in the absence of its cognate permease, which may be the first functional step in assembly of the transporter [99].

Vanadate-catalyzed photocleavage of the ABC of this protein recently demonstrated that the signature sequence LSGGQ from the first monomer is adjacent to the Walker A motif of the second monomer, and both motifs are directly involved in ATP binding and hydrolysis [58]. ATP binds cooperatively to the ABCs of two monomeric NBDs and this is believed to initiate the formation of a 'nucleotide sandwich dimer', as was found with Rad50 [83], and MJ0796 of *M. jannaschii* [143, 193]. It must be noted, however, that a mutant MJ0796 was used that cannot hydrolyze ATP [143, 193], and therefore may not be the most suitable tool to investigate ATP-dependent dimerization. In the HisP dimer the signature sequence and the Walker A motif of both monomers are too far apart to be functionally interacting [85], and it has by now been disputed whether here the functional dimer was crystallized. MalK of *T. litoralis* was crystallized as a dimer, but differs from more recently determined structures as it deviates from two-fold symmetry at the dimer interface [50].

The close proximity of the ABC motifs from opposing halves of the molecule has also been found in P-glycoprotein, indicating that this is a functional phenomenon not only in soluble NBDs [125]. It is difficult to reconcile this model for ATP-binding with the MsbA structure, where the NBDs are spatially far apart [104]. On the other hand, application to the BtuCD structure [122] suggests that the conformational changes associated with the dimerization of the NBDs compose a critical step in substrate translocation [193]. Thus, the 'nucleotide sandwich dimer' model for ATP-driven functional dimerization of the NBDs is gaining more support [83, 93, 193].

## 5. Transport mechanism of binding protein-dependent transporters

The exact mechanism of binding protein-dependent ABC transporters is unknown. Nevertheless, a substantial amount of data collected over the years on the respective subunits of these complex transporters, as well as on related proteins, has been combined into a model. In the absence of sufficient biochemical and structural information regarding the functional interactions in binding protein-dependent secondary transporters, an attempt is made here to derive a preliminary model based on data obtained with both secondary and binding protein-dependent primary transporters.

*5.1 Substrate binding*

In 1982, the theoretical 'Venus'-fly-trap model' was postulated, which proposes a uniform mode of interaction of extramembranal solute-binding proteins with their respective substrates [131]. According to this model, the substrate is bound in the cleft that separates both globular domains, followed by 'closure' of the protein. The movement of both globular domains has been demonstrated in studies on the histidine-binding protein HisJ of *S. typhimurium* [109, 237, 238], the ribose-binding RBP [21], and the D-allose-binding ALBP [128] of *E. coli*. The vitamin $B_{12}$-binding BtuF protein of *E. coli* has a more rigid $\alpha$-helical 'hinge' region (see section 4.2), therefore the movement of both lobes upon substrate binding is less pronounced, although still observed [27, 97]. Nevertheless, the chemical properties of the wide range of structurally unrelated potential substrates of this class of proteins require flexibility in substrate binding at the molecular level [7, 77, 200].

In all cases investigated thus far, extensive hydrogen bonding of the substrate to residues in the substrate-binding cleft plays a crucial role [170]. For example, amino acids such as histidine and glutamine are specifically bound via hydrogen bonding by HisJ of *S. typhimurium* [154] and GlnBP of *E. coli* [84], respectively. Charged, polar, and non-polar residues, as well as the peptide backbone of the protein are involved in these interactions and stabilize opposing charges of ligand and protein [84, 154]. Interestingly, carbohydrates are bound via hydrogen bonding as well as stacking of the hydrophobic regions of the sugar rings against aromatic residues from the peptide backbone, which stabilizes binding [170]. This is the case with both MBP [10] and ALBP [35] of *E. coli*. The trehalose/maltose-binding protein TMBP of the hyperthermophilic archaeon *T. litoralis* is structurally homologous to MBP, although the amino acid sequence identity is only 28% [51]. Strikingly, the number of hydrogen bonds between TMBP and its substrate is significantly higher than that of MBP, i.e. 14 compared to 5 with the first, equivalent, glucose moiety of the bound substrate [51]. Thus, despite high structural similarity, the actual binding mode of these proteins can differ significantly.

Strikingly, the electrostatic surface potential of the substrate-binding cleft does not affect binding of charged substrates, as has been shown both for the phosphate-binding

PBP and sulfate-binding SBP proteins of *E. coli* [139]. Upon binding, the protein engulfs the phosphate molecule as described above and binding is stabilized by 12 hydrogen bonds. Although the electrostatic potential of the phosphate-binding region of PBP is strongly negative, the protein specifically binds its negatively charged substrate [139]. The same holds true for SBP.

Despite their high specificity, several solute-binding proteins are able to bind more than one substrate. The maltose ABC transporter of the archaeon *Pyrococcus furiosus* has been shown also to transport trehalose, whereas the maltodextrin transport system of this organism transports maltotriose and higher malto-oligosaccharides [106]. MBP of *E. coli* binds maltose, a disaccharide, but also a range of maltodextrins up to maltoheptaose, and derivatives thereof [26]. However, not all of these substrates are transported into the cytoplasm. UV-spectral analysis of binding of non-transported compounds revealed that these are all bound by MBP in a manner that results in a blue shift in its UV spectrum (hence referred to as B-mode), whereas transported substrates such as maltose cause a red-shift in its UV spectrum (R-mode) [71]. This indicates that the latter mode is the physiologically active one. Indeed, electron paramagnetic resonance studies on binding of substrates by MBP revealed that in the 'B-mode' of binding, the substrate prevents closure of the two domains of MBP [72]. As a result, the MBP-ligand complex is unable to functionally interact with its cognate permease [72] (see section 5.2). These findings indicate that binding of a solute by the extramembranal component does not automatically mean that the compound is transported by the permease.

Sequence-unspecific oligopeptide binding by $OppA^{Sty}$ of *S. typhimurium* is mediated via hydrogen bonding of the peptide backbone to $OppA^{Sty}$ [202]. Peptides consisting of two up to five amino acid residues are bound according to the Venus' fly-trap model which would normally impose tight specificity on binding. However, the peptide side chains are accommodated in water-filled cavities [202]. Intriguingly, a similar mode of oligopeptide-binding has been proposed for the membrane domains of the binding-protein-independent ABC transporter TAP [208]. It must be noted, however, that for this transporter both the N- and C-terminal residues of the model nonapeptide are important for binding [208], whereas this is not the case for $OppA^{Sty}$ [202]. In contrast, $OppA^{Lla}$ from *L. lactis* exhibits different binding properties, preferring larger substrates than the

functional homolog from *S. typhimurium*. Here, not the termini of the oligopeptide, but the central amino acid residues at positions 4, 5, and 6 are crucial for substrate binding [203]. In addition, $OppA^{Lla}$ only binds residues 1-6, the remainder of the peptide is proposed to stick out of the binding cleft [203].

To date, only one member of the binding protein-dependent secondary transporters, i.e. the $C_4$-dicarboxylate-binding protein DctP of *Rh. capsulatus*, has been studied in detail. Conformational changes in the presence and absence of substrate were probed, using stopped-flow fluorescence techniques as well as steady state fluorescence spectroscopy [231, 232]. These experiments revealed that the protein binds its substrates with unequal affinity: L-malate > succinate > fumarate > D-malate (Fig. 3a) [232]. The binding of substrate affects the equilibrium between open and closed conformations of the protein, accelerating the transition between conformations [231]. In addition, a salt bridge appears to be involved in binding of the anionic substrate, playing a role in stabilization of the closed conformation [231].

*5.2 Functional interactions in the binding protein-dependent ABC transporter*

Following substrate binding, the extramembranal substrate-binding protein docks onto the extracellular side of the integral membrane permease. Both lobes of the substrate-binding domain interact with the permease, as has been demonstrated for RBP of *E. coli* [158], HisJ of *S. typhimurium* [120], and BtuF of *E. coli* [27]. The ribose transporter permease RbsC is present as a homodimer in the complete transporter, and RBP interacts symmetrically with both monomers in the permease [158]. As described above, MBP can only interact effectively with its cognate permease domain if the bound substrate allows sufficient closure of both lobes [71, 72]. Otherwise, MBP can not associate with the permease in the correct manner required for ATPase stimulation and subsequent release of the substrate [70]. Two negatively charged 'knobs' on the lobes of BtuF interact with two positively charged 'pockets' on the periplasmic surface of the permease domain BtuC, positioning the bound $B_{12}$ molecule directly over the translocation path [27]. The residues that are involved may form salt bridges that would account for the stability of this interaction [27].

The maltose transporter of *E. coli* does not hydrolyse ATP in the absence of MBP [45]. In contrast, the histidine transporter displays a low intrinsic ATPase activity [119]. Binding of the extramembranal domain to the permease domain stimulates ATPase activity in the cytoplasmic NBDs [9, 36, 119, 120]. Interestingly, unliganded MBP is able to bind to and stimulate the ATPase activity of a mutant $MalFGK_2$ permease that subsequently is capable of transporting lactose despite the fact that MBP does not bind this sugar [140]. These results indicate that the membrane domains of the transporter also play a part in determining the substrate specificity of the system [46, 140]. Both ligand-free and ligand-bound HisJ bind to the histidine permease, but the ligand-bound protein more efficiently stimulates the ATPase activity, while unliganded HisJ does not compete with liganded HisJ for the stimulation of the ATPase activity of the permease [9, 119]. In this regard, the membrane domains HisQ and HisM of the histidine transporter seem to regulate the ATPase activity of HisP in the absence of HisJ, as this activity is substantially higher in the absence of the permease domains [119, 150]. These findings show that apart from high-affinity binding of and delivering of the substrate to the permease, the extramembranal substrate-binding protein plays an important role in activating the transporter [46].

In order to be able to detect the signal that is provided by docking of the substrate-binding protein, the NBDs should be in close contact with the membrane domains of the transporter. Indeed, the conserved EAA-region of binding protein-dependent ABC transporters [186] has been shown to interact with the NBDs of the maltose transporter of *E. coli* [144]. Moreover, in the recently published three-dimensional structure of BtuCD this region is found in the L-loop that clearly interacts with the NBD [122]. A similar loop has been found to be important in the membrane domain-NBD interactions of CFTR, and is present in the first cytoplasmic loop of a range of ABC drug-exporters such as P-gp and LmrA [122]. The interaction between the permease and NBDs of MsbA takes place via two substantially larger intracellular domains, i.e. ICD1 and ICD 2, which are conserved in MDR-like ABC transporters [34]. The 'L-loop' as defined for BtuC is found in the ICD1 of MsbA, supporting the view that this region is important in these transport systems.

Differing views exist on how the NBDs associate with the permease. The MalK protein was found to dimerize already in the absence of its cognate permease [99],

whereas the two HisP molecules of the histidine transporter are recruited independently to the membrane domains of the transporter [121]. In addition, although the ABC type ATPases show positive cooperativity in ATP-hydrolysis when present in the permease [44, 119], or as soluble domains [143], the purified HisP ATPase behaves differently [150]. The positive cooperativity of the ATPase dimer is explained by the finding that the ATP-binding cassettes from opposite monomers have to interact in ATP binding in order to achieve ATP-hydrolysis (see section 4.4) [58, 125, 143, 193]. Following dimer formation, the $\alpha$-helical domain of the ABC is thought to interact with the hydrophobic membrane subunits of the transporter [188]. In the BtuCD structure, residues present in the Q-loop of the NBD BtuD contact the membrane domain BtuC in its L-loop [122]. Mutations in these and anologous residues in different ABC transporters, such as CFTR and TAP1, strongly affect the transport activity [122]. Thus, the contact regions described above are remarkably conserved within the ABC transport protein family, although not always at identical relative positions within the molecule.

*5.3 Binding protein-dependent primary transport*

The information described above has allowed the construction of a fairly detailed model for binding protein-dependent ABC transport [43, 46]. First, the ligand-bound extramembranal domain tightly binds to the permease, positioning its bound substrate over a 'chamber' or 'cavity' in the membrane domain that is closed on the cytoplasmic side of the membrane [27, 122]. The tight association of the binding protein with the permease reduces its affinity for the ligand [36], resulting in its release into the receiving membrane domain. Subsequently, the substrate-binding site of the membrane domain has to reorient somehow to the cytosol, in order to facilitate it's release into the cytosol [36, 46]. These transporters are thought to possess one high-affinity substrate-binding site in the extramembranal domain, and one low-affinity substrate binding site in the membrane domain [46, 140], possibly explaining the directionality of the system. As these systems are inactive in the absence of the substrate-binding domain, the substrate-binding site of the membrane domain probably is occluded prior to docking of the extramembranal domain [140]. It must be noted, however, that although the BtuCD structure was obtained in the absence of BtuF, the membrane domains do form a

chamber that is accessible from the 'outside' of the membrane, and sufficiently large to fit the vitamin $B_{12}$ molecule [122]. The extacytoplasmic domain is still associated with the transporter after release of its ligand [36], possibly serving as a 'plug' that prevents the solute from escaping back into the periplasm [43, 46]. This domain may also impose directionality on the transport process.

The current model places one extramembranal substrate-binding domain at the external face of each permease. The ABC transporters with one or even two substrate-binding domain(s) fused to each membrane domain effectively will have two to four substrate binding domains present in each complex, as these systems are thought to form homodimers. This possibly enhances the efficiency of these systems [213]. Alternatively, they may form a functional binding site only upon dimerization. In this regard, the BtuF protein interacts as a monomer with its cognate permease [27].

Several signals are sent and received during binding protein-dependent transport. These are most likely transmitted via small conformational changes in the transmembrane $\alpha$-helices of the membrane domain resulting from binding of the substrate(-binding protein) [43, 46, 77, 78]. The close contacts between membrane domain and NBDs in the immediate vicinity of the ATP-binding cassette provide a pathway for the signal(s), possibly by reorienting the signature sequence of one NBD so that it interacts with an ATP molecule bound to the opposite NBD [43]. Dimerization may result in movement of connected transmembrane helices, causing the reorienting of those helices that are lining the translocation path, and thereby allowing the passage of the substrate [43, 46, 122]. Withdrawal of the signature sequence subsequently allows for the release of ADP and inorganic phosphate ($P_i$) [43].

The details of the effects of ATP binding and hydrolysis on the binding protein-dependent transport process await further investigation. However, studies addressing these questions have been performed with ABC-type export proteins. ATP binding and hydrolysis by LmrA results in significant changes both in the secondary and tertiary structure of the protein [225], caused by extensive 'cross-talk' between the membrane and nucleotide-binding domains [224]. The protein has been proposed to function as a two-cylinder engine, with two drug-binding sites with alternating affinity: high at the cytoplasmic side of the membrane, and low when facing the external side [217]. The conformational changes required for this transport mechanism are induced by cycles of

ATP binding and hydrolysis [217]. Although transport by the binding protein-dependent ABC transporters takes place in the opposite direction, the presence of two analogous substrate-binding sites has been suggested as described above [46], although such an organization of binding sites cannot be inferred from the current Btu structure.

Interestingly, binding of ATP already induces conformational changes in LmrA that are important for the release of the bound substrate [224], a phenomenon also observed in the homologous P-glycoprotein [126, 179]. Similar allosteric cross-talk between membrane domains and NBDs has been reported for TAP [63]. Although the NBDs of P-gp [179] and LmrA [217] are thought to function cooperatively, the NBDs in the functional TAP1-TAP2 heterodimer are thought to have distinct functions: ATP binding to the TAP2-NBD controls substrate binding and release, whereas the TAP1-NBD controls the structural rearrangements of the transmembrane substrate pathway [215]. Nevertheless, the observed conformational changes in one domain of the transporter that are invoked by movements of (part(s) of) a second, connected domain, can also be envisaged in transporters where these domains are expressed as separate polypeptides, as these subunits associate via the conserved tight contacts described above.

Thus far, the exact stoichiometry of hydrolyzed ATP per transported substrate molecule has not been firmly established. Reported numbers range from 1-2 ATP molecules per transported substrate [119], to 5 ATP molecules per 2-3 transported peptides [215].

### *5.4 Binding protein-dependent secondary transport*

As was noted by Higgins in 1992 [77], in the absence of structural information on the nature of a transporter, our understanding of the molecular basis of solute transport is superficial. Nevertheless, constructing a model that incorporates whatever information is available is worth considering, if only to eliminate certain possibilities and provide others to be excluded. Currently, no structural information is available on the members of the binding protein-dependent secondary transporter family. Nevertheless, we can attempt to model the substrate translocation process of these systems on the information that is available on both secondary and binding protein-dependent transporters. One important aspect to keep in mind, though, is the lack of significant

amino acid sequence homology of any of the binding protein-dependent secondary transporters with proteins outside the family [53, 98, 171]. In particular, the small 4-TMD integral membrane subunit has no counterparts in any known membrane protein family.

It is known for the substrate-binding domains of ABC transporters that they form remarkably similar structures in the absence of significant amino acid sequence identity [7, 62, 170, 200]. Therefore it may be reasonable to assume that the extramembranal substrate-binding proteins of secondary transporters bind their substrates according to the Venus' fly-trap model (see section 5.1). This will result in closure of two globular domains, bringing them together and thereby creating a binding site on the protein with high affinity for the integral membrane permease. For convenience, the stoichiometry of one substrate-binding protein per integral membrane permease will be assumed in the following steps.

Docking of the substrate-binding domain will result in a conformational change in the permease, which reorients the substrate-binding region within the transmembrane segments so that it can accept the substrate. The requirement for this step can be deduced from the fact that no substrate translocation takes place in the absence of the glutamate-binding protein of *R. sphaeroides* [91], which means that the substrate-accepting site of the permease probably is occluded in its absence. The presence of one substrate-accepting region in each permease is assumed based on data published for secondary transporters that, although sometimes functional as a cooperative dimer, contain a single substrate translocation pathway per monomer [94, 220].

In analogy to the mechanism described for the binding protein-dependent ABC transporters (section 5.3) [27, 43, 158], this docking will require the presence of a more or less symmetrical permease domain. However, the permease is made up of two differently sized subunits (see section 2, Fig. 1b), and whether these are present as mono- or multimers is unknown at this time. A sketch of some of the possible quaternary conformations that may be assumed by the membrane subunits is presented in Fig. 6. One possibility assumes that the large subunit is present as a symmetrical 'OxlT-like' [81] molecule, i.e. with the putative substrate translocation pathway located in between both halves of the monomer (Fig. 6a).

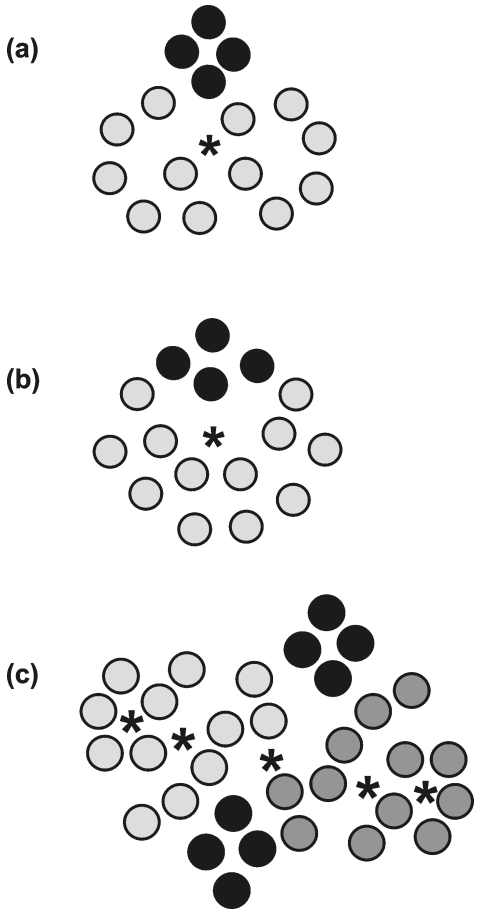



**Figure 6.** Sketchbook representation ('top view') of three possible structural models of a binding protein-dependent secondary transporter. (a) Monomeric organisation: the large domain is present as an 'OxlT-like' symmetrical monomer, with one putative substrate translocation path present at the interface between both halves of the large permease subunit. (b) Monomeric organisation: the small permease subunit is an integral part of the substrate translocation pathway. (c) Dimeric organisation: the large domain is drawn as a 'NhaA/MelB-like' dimer with two putative substrate translocation pathways within each monomer, and an additional possible pathway at the dimer interface. In all cases, the small domain provides the main docking site for the substrate-binding protein. Transmembrane helices are represented as circles. Black: 4-TMD subunit; greys: 12-TMD subunit. Asterisk: putative substrate translocation pathway.

The small membrane domain provides the interaction site for the substrate-binding domain, possibly via two periplasmic loops [241]. We cannot exclude the possibility that the substrate-binding domain interacts with both membrane proteins, in order to position the substrate over the substrate translocation pathway as described in the Btu-model [27]. In a second model, the small membrane domain not only provides this

docking site, but is an integral part of the substrate translocation pathway (Fig. 6b). Preliminary experiments show that both proteins are required to achieve expression of the membrane domains of the YiaMNO transporter, indicating that both subunits stabilize each other [Chapter 2]. This could mean that they are in close contact within the functional permease. Finally, in a multimeric structure the large membrane domain is depicted as a 'NhaA-like' [235] monomer, that resembles the MelB structure [69], in a dimeric complex. Possible substrate translocation pathways are indicated within each monomeric half of the complex, but may also be present at the dimer interface (Fig. 6c).

Unidirectional transport by these systems is energized by the pmf and/or smf [53, 98, 171] which basically would resemble the coupling ion-symport mechanism (see section 2.1) once the substrate has been transferred from the substrate-binding domain into the substrate translocation pathway of the permease. A molecular mechanism for this transport mode has been modelled on the (monomeric) lactose permease LacY of *E. coli* [94]. Residues have been identified that are essential for coupling, and these are located in close proximity to the substrate translocation pathway. Thus, conformational changes resulting from substrate binding facilitate binding of the coupling proton. After release of the ligand into the cytosol, the protein must return to its ground state. The proton is most likely released into a water-filled crevice in between helices, and acted upon by both the membrane potential (negative inside) and the transmembrane proton gradient [94]. Future biochemical studies of these transporters, in conjuction with structural elucidation, should reveal the mechanism of function.

## 6. Evolution of transport protein families

The structural organisation of binding protein-dependent secondary transporters raises questions about their evolutionary origin [53, 98, 171]. These systems could easily be perceived as constituting a 'mix' of the secondary and the binding protein-dependent primary transporters. But, did they arise from or precede these systems? Or, alternatively, did they evolve in parallel? In recent years, the evolution of transport systems is viewed more in the context of acquisition and loss of (parts of) structural domains, than in amino acid sequence homology alone [54, 180]. In this regard, binding protein-dependent secondary and primary transporters have been identified in both

prokaryotic Kingdoms of life [53, 98, 171], whereas the secondary transporters are found in all three Kingdoms [163]. Although this means that all three transport protein families originated before the separation of the Bacterial and Archaeal branches, it indicates that the latter may be older. It has therefore been suggested that the secondary transporters in some way resemble the 'primordial' transporter: a single membrane protein catalyzing (bi-directional) transport activity. During evolution, an extramembranal substrate-binding domain has associated with a secondary transporter, thereby conferring high affinity as well as directionality on the system. The additional membrane domain may have become essential as the putative docking site for the extramembranal domain [54]. This small domain domain may even play an additional role, as it appears essential for maintaining stability of the transporter [Chapter 2]. Interestingly, bacterial arsenate (Ars) transporters are fully functional as secondary transporters. However, they are also able to associate with an ATPase domain, which radically changes their mode of energization as well as improving the transport efficiency of the system [177]. One step further would mean the association of an extracellular substrate-binding domain to the system, creating a binding protein-dependent primary transporter [54]. The successful spreading of these systems throughout a range of unrelated prokaryotes may have taken place via horizontal gene transfer [107], as systems that transport identical or similar compounds are more homologous than systems that originate from the same organism [40, 110, 185, 186, 200]. The conservation of gene order in operons that specify similar functions, such as these homologous binding protein-dependent transporters, indicates that the selective processes that are involved maintain this organisation [201, 239].

Stikingly, a recent genomic survey revealed that the clustering of three enzymes that function in carbohydrate metabolism is coupled to the presence of a putative transport system, although the class of transporter differs [Chapter 5]. In addition, the ability to phosphorylate the carbohydrate substrate has been conserved, either in the nature of the transporter, i.e. a PTS system, or the presence of a gene encoding a kinase. Evolution, it seems, can be an extremely subtle process, as in one case only the transporter at the centre of the operon has been exchanged for a protein from different class, leaving the remainder of the operon in tact [Chapter 5].

## 7. Concluding remarks

Transport proteins are recognized as essential, as they confer a clear selective advantage [159]. The binding protein-dependent secondary transporters are an ancient transport protein family, which has been discovered only recently [53, 98, 171]. In order to gain insight into the function of these systems as well as their transport mechanism, more members need to be characterized. Pending the availability of detailed biochemical and structural information on this intriguing class of transporters, combining similar data that has been obtained with the 'classical' secondary and the binding protein-dependent primary transporters has given us some room for speculation. Whether the model described here has any truth to it will be revealed in the upcoming years.

## 8. Scope of this thesis

This study is concerned with the function and molecular mechanism of the binding protein-dependent secondary transporter of *E. coli* K-12 encoded by the *yiaMNO* genes. First, these genes were cloned and overexpressed [Chapter 2]. Preliminary expression data indicate that both membrane domains stabilize each other. In addition, the periplasmic substrate-binding domain YiaO can be readily overexpressed and purified. Second, an unmarked genomic deletion mutant of the *yiaMNO* genes was constructed in strain MC4100 to study the physiological role of the transporter [Chapter 3]. Several intriguing, but as of yet unexplained physiological effects were observed in the deletion mutant. Constructing the *yiaMNO* deletion in a strain that constitutively expresses the *yiaKLMNOPQRS* gene cluster resulted in the identification of the rare pentose L-xylulose as a substrate for the transporter [Chapter 4]. Finally, a genomic survey revealed that the clustering of the *yiaQRS* genes is highly conserved throughout a range of Gram-positive and Gram-negative bacteria, most of which are known pathogens of man [Chapter 5]. Strikingly, in all cases these gene clusters also contain a transporter, but the class of transporter varies. The possible fate and function of L-xylulose following uptake by the YiaMNO transporter are discussed [Chapter 6].

# *Chapter 2*

## Overexpression and preliminary characterization of the *yiaMNO* genes of *Escherichia coli* K-12

*"When you are trying to solve a problem,*
*it is always convenient to know the answer."* [23]

Titia H. Plantinga, Chris van der Does, Wil N. Konings and Arnold J.M. Driessen

**Summary**

The *yiaMNO* genes of the *Escherichia coli* K-12 genome encode a member of the recently defined family of binding protein-dependent secondary or tripartite ATP-independent periplasmic (TRAP) transporters. This chapter describes our efforts to obtain an overexpression system for the YiaMNO transporter. The membrane proteins YiaM and YiaN could only be expressed from a synthetic operon that contained at least both the *yiaM* and the *yiaN* genes, suggesting that both proteins stabilize each other. The periplasmic substrate-binding protein YiaO could readily be overexpressed in and purified from the periplasm of *E. coli*, and was processed at the consensus Gram-negative signal sequence cleavage site.

## Introduction

Binding protein-dependent secondary [53, 54], or tri-partite ATP-independent periplasmic (TRAP) [59, 98, 171] transporters mediate the uptake of solutes in a binding protein-dependent manner. Transport is driven by the proton-motive force (pmf) instead of ATP-hydrolysis, which distinguishes these transporters from the binding protein-dependent members of the ATP-binding cassette (ABC) superfamily of transport proteins [46, 78, 79, 86]. The solute-binding protein donates the substrate to the membrane domain of the transporter, which consists of two dissimilar protein subunits in unknown stoichiometry. The large subunit consists of 12 putative transmembrane domains (TMDs) and one large, cytosolic loop separating TMD 6 from TMD 7. This architecture resembles that of 'classical' secondary transporters. The small subunit consists of 4 putative TMDs [53, 171, 241]. In some cases the membrane domains are 'fused' into one large polypeptide.

Both the mechanism of and the subunit interactions involved in transport by multi-component uptake systems have been a major topic of research in the past years. Most studies have focused on the most easily accessible components of these transporters, i.e. the solute-binding protein [5, 35, 70, 71, 72, 128, 154, 207, 237, 238] or the cytosolically localized nucleotide-binding domains (NBDs) [58, 99, 193, 232] of ABC transporters. Additional reports address the interaction of either the solute-binding protein [9, 27, 119, 158], or the NBDs [121, 144] with the membrane domains.

There have been only a few reports in which all subunits of a multi-component ABC transporter have been overexpressed, purified and functionally reconstituted. This type of experiment has been used to show that transport by both the maltose transporter of *Escherichia coli* and the histidine permease of *Salmonella typhimurium* strictly depends on ATP-hydrolysis [47, 118]. Up to now, only reports exist of studies performed with isolated components of binding protein-dependent secondary transporters. The glutamate-binding protein of *Rhodobacter sphaeroides* was isolated from the periplasmic fraction and shown to function together with a secondary membrane permease in an ATP-independent manner [91]. The solute-binding protein DctP of the DctPQM $C_4$-dicarboxylate transporter of *Rhodobacter capsulatus* was expressed in the periplasm and N-terminally processed at the predicted signal sequence cleavage site [190], as was the

ectoine-binding protein TeaC of the TeaABC transporter of *Halomonas elongata* when heterologously overexpressed in the periplasm of *E. coli* [206].

A purified and reconstituted transporter would allow the detailed study of transport kinetics, as well as shed light on the mode of interaction and the roles of the respective subunits. Here, we describe our efforts to overexpress the three protein subunits of YiaMNO, the binding protein-dependent secondary transporter of *E. coli* K-12. Expression of YiaM, the small membrane subunit, required co-expression of at least YiaN, the large subunit, which indicates that both proteins stabilize each other. YiaO, the periplasmic solute-binding protein, was readily overexpressed in its native form, and N-terminal sequencing showed the protein is processed at the predicted signal sequence cleavage site. These studies are a first step towards a functionally reconstituted system.

## Experimental procedures

*Bacterial strains and growth conditions*

All strains used in this study are *E. coli* K-12 derivatives: DH5α (*supE*44, *hsdR*17, *rec*A1, *end*A1, *gyrA*96, *thi*-1, *rel*A1, Δ*lacU*169 (*80 *lacZ*ΔM15); laboratory collection) was used for cloning, and SF100 (KS272 {F⁻ ΔlacX74 *galE galK thi rpsL*(*strA*) Δ*phoA*(*PvuII*)}, ΔOmpT) [18] was used for overexpression studies. Strains were grown aerobically in Luria-Bertani (LB) medium [141] at 37°C. Ampicillin was used at 50 μg/ml final concentration.

**Table 1.** Primers used in this study

| Gene: | Primer: | Sequentie: | Site[c]: |
|---|---|---|---|
| *yiaM* | p1 (f)[a] | TAGCCATGGAAAAAATACTCGAAGC | *Nco*I |
| *yiaM* | p2 (r)[b] | GGGGAATTCTTAAGCTCCTTGCGG | *Eco*RI |
| *yiaN* | p3 (f) | TAAGAATTCAGGAGAGTAATCATGGCTGTGCTG | *Eco*RI |
| *yiaN* | p4 (r) | GCGGGATCCTTAATCCATTTCAAAGGG | *Bam*HI |
| *yiaO* | p5 (f) | AATGGATCCATTAAAAGGAAAATATTATG | *Bam*HI |
| *yiaO* | p6 (r) | CCCTCTAGATTATTGCACCTCATCCAC | *Xba*I |
| *yiaO* | p7 (f) | GGGCCATGGAATTACGCTCTGTAACC | *Nco*I |
| *yiaN* | p8 (f) | TTAACCATGGCTGTGCTGATT | *Nco*I |
| *yiaO* | p10 (r) | CCCGGATCCACCTTGCACCTCATCCAC | *Bam*HI |

[a] forward primer
[b] reverse primer
[c] restriction endonuclease site introduced by primer

*Cloning of the* yiaMNO *genes*

All genes were cloned via PCR with the primers listed in Table 1, and primer positions are indicated in Fig. 1A. An overview of all vectors is given in Table 2 and Fig. 1B. The *yiaM* gene was ligated into vector pET302, resulting in vector pET910, introducing an N-terminal 6xHis-tag. The *yiaN* gene was cloned in two ways: either with primer pair p3/p4 for ligation in tandem with *yiaM* (vectors pET909 and pET900), or with primer pair p8/p4 for cloning into pET302, resulting in vector pET912 and introducing an N-terminal 6xHis-tag. The *yiaO* gene was cloned in three different ways: primer pair p5/p6 for cloning in tandem with *yiaMN* in pET900, primer pair p7/p10 for cloning into pET302, resulting in pET917 as described [Chapter 4], and primer pair p7/p6 for cloning into pET324, resulting in pET925.

**Table 2.** Vectors used in this study

| Vector: | Relevant characteristics: | Reference: |
|---|---|---|
| pET302 | Expression vector, *trc* promoter, $Am^r$, N-terminal 6x-His coding region | 211 |
| pET324 | Expression vector, *trc* promoter, $Am^r$ | 211 |
| pSA5 | Expression vector, *trc* promoter, $Am^r$, C-terminal 6x-His coding region | 6 |
| pET900 | pET302 carrying the *yiaMNO* genes | This work |
| pET909 | pET302 carrying the *yiaMN* genes | This work |
| pET910 | pET302 carrying the *yiaM* gene | This work |
| pET912 | pET302 carrying the *yiaN* gene | This work |
| pET917 | pSA5 carrying the *yiaO* gene | Chapter 4 |
| pET925 | pET324 carrying the *yiaO* gene | This work |

*Expression of the YiaMNO proteins*

SF100 cells carrying the pET302, pET910, pET912, pET909, or pET900 vector were grown in 100 ml cultures to an $OD_{660}$ of 0.8. Expression was induced by addition of 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). After 2 h of induction, the cells were harvested (4,000 x g, 10 min, 4ºC), resuspended in lysis buffer (10 mM sodium phosphate, pH 7.0; 30 mM NaCl; 0.25% (v/v) Tween 20; 10 mM ethylene dinitrilo tetraacetic acid (EDTA); 10 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA)), and sonicated (15 sec on/15 sec off, 10 cycles, 6 μ amplitude) using a Soniprep 150 sonicator (Beun de Ronde, Abcoude, The Netherlands). NaCl was added to 0.5 M final concentration, and the samples were centrifuged (12,000 x g, 20 min, 4°C). The supernatant of this step ('crude extract') and the pellet ('crude membranes') were analyzed as described below. Expression of His-tagged proteins was analyzed via semi-dry Western Blotting (Biorad, Veenendaal, The Netherlands) of the samples onto polyvinylidene fluoride (PVDF) membranes in blotting buffer (40 mM Tris; 40 mM glycine;

20% (v/v) methanol). All washing and incubation steps were performed in PBST (phosphate buffered saline containing 0.1% Tween 20) containing 0.1% (w/v) I-block (Tropix, Bedford MA, USA). The His-tag was detected using anti-Histidine-tag antibody (Dianova, Hamburg, Germany) followed by anti-mouse-antibody (Sigma-Aldrich Chemie, Steinheim, Germany). The signal was detected using the nitroblock/CSPD chemiluminescence-detection system (Tropix, Bedford MA, USA).
For overexpression of YiaO, cells transformed with vectors pET302, pET900, pET909, pET324, pET925, pSA5 or pET917, were induced with IPTG as described above. After 2 h of induction, cells were harvested (4,000 x g, 10 min, 4ºC), and the periplasmic fraction was isolated as described elsewhere [148].

*Purification of the periplasmic YiaO protein*

Native YiaO was overexpressed from pET925 as described above. For cation-exchange chromatography, the periplasmic samples were buffered with 50 mM malonic acid, pH 4.0-5.0 (buffer A), and incubated in batch for 45 minutes at 4ºC with SP-sepharose slurry (Amersham Pharmacia Biotech, Piscataway NJ, USA). The flow-through of each sample was collected, the column material was washed once with buffer A, and the sample was eluted with buffer A containing 1 M NaCl.
For anion-exchange chromatography, the periplasmic fraction was buffered with 50 mM Tris-HCl, pH 7.0-9.0 (buffer B). The samples were loaded onto a MonoQ HR 10/10 column connected to a fast performance liquid chromatography (FPLC) system (Amersham Pharmacia Biotech, Piscataway NJ, USA) at a flowrate of 0.5 ml/min. Proteins were eluted using a gradient of 0-200 mM NaCl in 20 min (1 ml/min) in buffer B.
For the size-exclusion experiments, the periplasmic fraction was buffered with 50 mM Tris-HCl, pH 7.5; 10 mM NaCl, at 4ºC. The protein was isolated via size-exclusion chromatography, using an HR 16/60 column connected to a GradiFrac System (Amersham Pharmacia Biotech, Piscataway NJ, USA) at a flow rate of 0.3 ml/min. Column fractions were analyzed as described below.
The C-terminally 6xHis-tagged YiaO protein (YiaO-His) was overexpressed from pET917 and purified via $Ni^{2+}$-NTA affinity chromatography as described elsewhere [Chapter 4].

*Iso-electric focusing*

MonoQ-fractions containing highly enriched YiaO were pooled and dialyzed overnight at 4°C against 10% (v/v) glycerol in demineralized water. The sample was applied in a Protean IEF (Isoelectric focusing) system (Biorad, Veenendaal, The Netherlands), with 0.1 M $H_3PO_4$ at the cathode and 0.1 M NaOH at the anode. The run was performed for 4 h at 3000 V/300 mA (12 W limit), at 4°C. Fractions were collected and analyzed.

*Miscellaneous methods*

The *E. coli* K-12 YiaO amino acid sequence was used as query to perform a standard BLAST (Basic Local Alignment Search Tool) search at http://www.ncbi.nlm.nih.gov/BLAST.
Purified YiaO was blotted onto a PVDF membrane and N-terminally sequenced by the Protein Service Laboratory of the University of British Colombia (Vancouver BC, Canada). All overexpressed proteins and purified fractions were analyzed on 12% sodium docecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by Coomassie Brilliant Blue (CBB) and silver staining. Protein content was determined using the Biorad $D_C$ method (Veenendaal , The Netherlands).

## Results

*Expression of YiaM is stabilized by co-expression of YiaN*

The genes encoding the membrane components of the *E. coli* YiaMNO transporter YiaM and YiaN were cloned separately into vectors that fused an N-terminal His-tag to the proteins (vectors pET910 and pET912, respectively, Fig. 1B). However, using an anti-His-tag antibody, no expression was detected in crude membrane or cytosolic extracts of SF100 cells transformed with vectors containing either the *his-yiaM* (pET910) or the *his-yiaN* (pET912) gene (data not shown). Nonetheless, when the genes were cloned in a synthetic operon in the original order, together with the periplasmic substrate-binding protein YiaO (vector pET900, Fig. 1B), analysis of the periplasmic fraction of these cells showed that YiaO was expressed (Fig. 2A, lane 2, compare to lane 1: 'empty vector' pET302). In this construct, the *yiaO* gene is third in order, suggesting that *yiaM* and *yiaN* were also expressed in these cells. When expressed from pET900, the *yiaM* gene is fused to an N-terminal His-tag. Indeed, when the crude membrane fraction of cells containing pET900 was analyzed using an anti-His-tag antibody, His-YiaM was detected as a single band (Fig. 2B, lane 5). The protein runs at a slightly higher molecular weight than the 17.5 kDa that is predicted from sequence analysis, which is probably due to the introduced His-tag. In addition, analysis of the crude membrane fraction of cells containing vector pET909 with the *His-yiaM* and *yiaN* genes in tandem (Fig. 1B) showed that His-YiaM is expressed from this construct (Fig. 2b, lane 6).

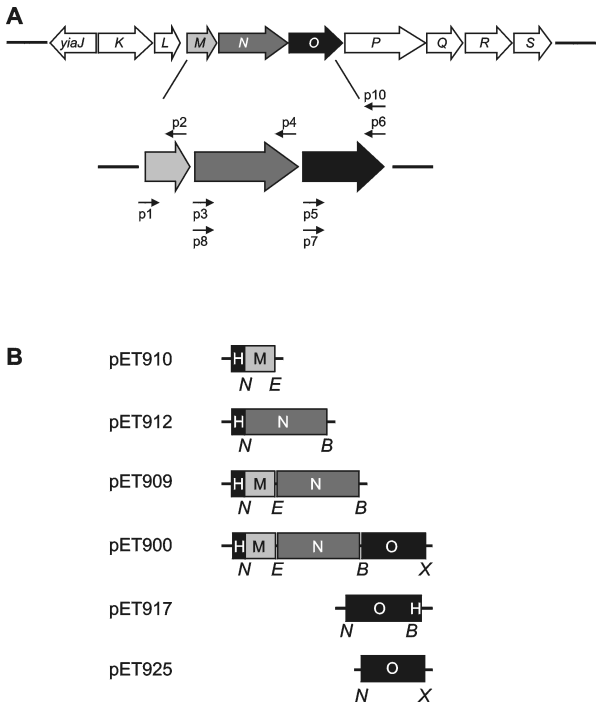



**Figure 1.** (A) Organization of the *yiaJ-S* gene cluster and location of the primers (listed in Table 1) used to clone the separate genes. (B) Schematic representation of the vectors used to express the *yiaMNO* genes, either as separate constructs or in tandem in artificial operons. Restriction sites are indicated as follows: *N*: *Nco*I; *E*: *Eco*RI; *B*: *Bam*HI; *X*: *Xba*I. 'H': 6xHis-tag-coding region.

Although detection of YiaN when expressed from this construct was not possible, these findings demonstrated that His-YiaM was expressed only when at least the *yiaN*-gene was co-expressed, indicating that both membrane proteins stabilize each other. The overexpression levels of both YiaM and YiaN were too low for detection by CBB or silver staining of SDS-PAGE gels of isolated membranes (data not shown).

*Overexpression and purification of the binding protein YiaO*

To perform substrate binding and transport experiments, sufficient amounts of the periplasmic component of the transporter, i.e. YiaO, need to be obtained. The protein was readily overexpressed in its native form in *E. coli* SF100 carrying pET925 (Fig. 3B). Several methods were used to purify the protein without making use of an introduced tag. The YiaO protein has a calculated pI of 6.1, which was confirmed by isoelectric focusing, showing that the protein has pI in the range of pH 5.5-6.5 (data not shown).

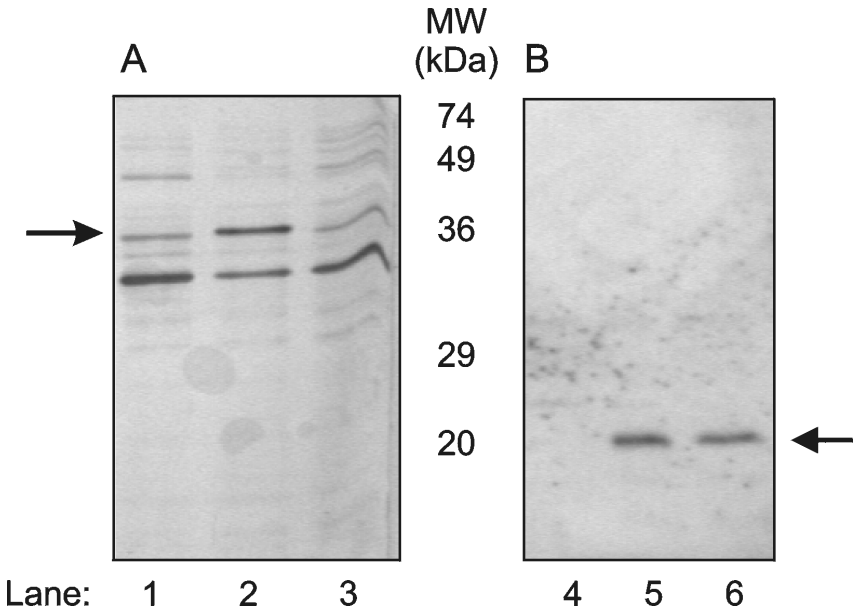



**Figure 2.** SDS-PAGE analysis of protein expression by SF100 cells carrying either pET302 (lanes 1 and 4), pET900 (lanes 2 and 5), or pET909 (lanes 3 and 6). (A) Periplasmic fractions, visualized by silver stain. Arrow: YiaO. (B) Western Blot analysis of crude membrane fractions. Detection: anti-His-tag antibody. Arrow: His-YiaM.

Purification by ion-exchange chromatography failed, as YiaO did not bind to the cation-exchange column material, and associated only very loosely with the anion-exchange column (data not shown). On the other hand, size-exclusion chromatography led to a significant purification of the protein (Fig. 3). The 36-kDa YiaO protein eluted in a single peak with minor contaminants, in particular the 30-kDa β-galactosidase introduced as the ampicillin-resistance marker (Fig. 3A). The identity of the YiaO protein (Fig. 3B, arrow) was confirmed by N-terminal amino acid sequencing (Fig. 4). The data showed that the protein is processed at the signal sequence typical for secreted proteins from Gram-negative bacteria (Fig. 4), as was predicted by the SignalP program [149].

YiaO was overexpressed in *E. coli* SF100 with a C-terminal 6xHis-tag, and purified to homogeneity from the periplasmic fraction by the cold osmotic shock procedure followed by $Ni^{2+}$-NTA-affinity chromatography (Fig. 5). This method proved to be more efficient than size-exclusion chromatography, and was therefore preferred in the experiments described in Chapter 4.

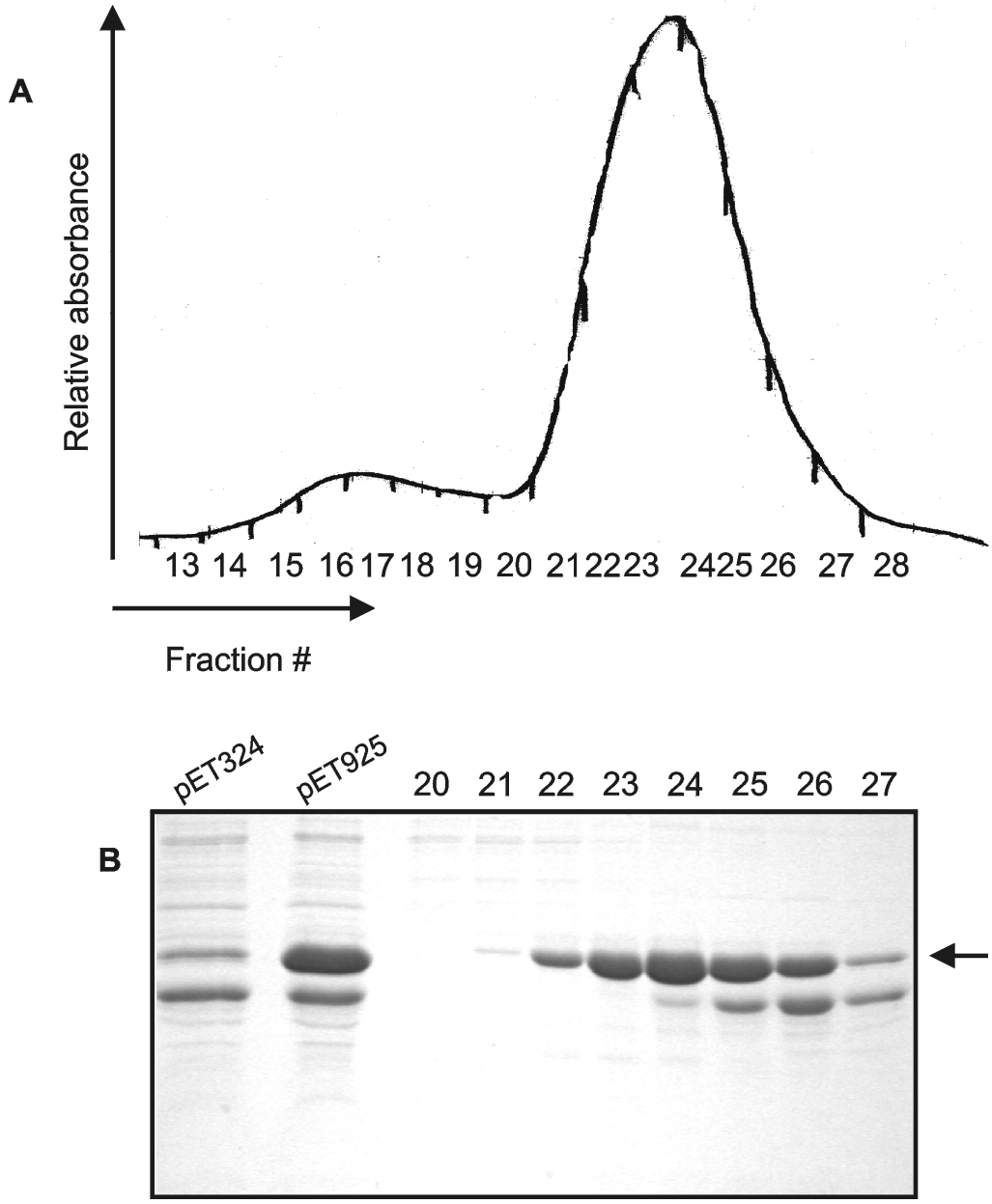



**Figure 3.** Overexpression and purification of native YiaO. (A) Size-exclusion chromatography of the periplasmic fraction of YiaO-overproducing cells. (B) SDS-PAGE analysis of periplasmic fractions of SF100 cells carrying the empty vector (pET324) or the YiaO-construct (pET925), and fractions 20-27 of the size-exclusion experiment. Proteins were visualized by silver stain. Arrow: YiaO.

```
YiaO_Ecoli      MKLR-SVTYALFIAGLAAFSTSSLAAQSLRFGYETSQTDSQHIAAKKFNDLLQERTKGEL 59
YiaO_Styph      MKLH-VIARSLLIAGLTVFSVSSLAAQSLRFGYETPQTDSQHIAAKKFNELLKEKTNGEL 59
YiaO_Hinfl      MKLFNFKKLSMLIAGFTLVTSPALAEISLRFGYEAPRSDSQHSAAKKFNDLLMKKTKGEI 60
PM1252_Pmult    MKRFNLKMLAALVAGMAVFTASANAATTLRFGYEAPRSDTQHEAAKKFNELLKEKTKGEI 60
                **       : ::**:: .:  .: *   :******:.::*:** ******:** ::*:**:

YiaO_Ecoli      KLKLFPDSTLGNAQAMISGVRGGTIDMEMSGSNNFAGLSPVMNLLDVPFLFRDTAHAHKT 119
YiaO_Styph      TLKLFPDSTLGNAQAMISGVRGGTIDMEMSGSNNFTGLAPVFNLLDVPFLFRDTAHAHKT 119
YiaO_Hinfl      KLKLFPDSTLGNAQTMISSVRGGTIDLEMSGSPNFTGLEPKLNVIDIPFIFKDREHVYKV 120
PM1252_Pmult    KLSLFPDSTLGNAQTMISAVRGGTIDLEMSGSPNFSGLVPKLNVIDIPFIFQNREHAYAV 120
                .*.***********:***.*******:***** **:** * :*::*:**:*::  *.: .

YiaO_Ecoli      LDGKVGDDLKASLEGKGLKVLAYWENGWRDVTNSRAPVKTPADLKGLKIRTNNSPMNIAA 179
YiaO_Styph      LDGKVGDELKKSLDSKGLKVLAYWENGWRDVTNSRAPVKTPGDLKGLKIRTNNSPMNIAA 179
YiaO_Hinfl      LDGEVGQNLLKDLEKQGLKGLAFWDVGFRAFSNSKQTVTKPEHIKGLKVRTNQNPMYIEA 180
PM1252_Pmult    LDGEIGQGLLKELEAQGLKGLAFWEVGFRSFTNSKHPVKTPDDIKGLKVRTNQNPMYIQA 180
                ***::*: *  .*: :*** **:*: *:* .:**: .*..* .:****:***:.** * *

YiaO_Ecoli      FKVFGANPIPMPFAEVYTGLETRTIDAQEHPINVVWSAKFFEVQKFLSLTHHAYSPLLVV 239
YiaO_Styph      FKIFGANPIPMPFSEVYTGLETRTIDAQEHPINVVWSAKFYEVQKYLSLTHHAYSPLLLV 239
YiaO_Hinfl      FKLLGSNPVPMPLAELYTALETRAVDAQEHPIGIFWSSKLYEVQKYLSLTNHGYTPLIVV 240
PM1252_Pmult    FSILGANPVPMPLSELYTALETRAVDAQEHPVGIVWSAKLYEVQKHLSLTNHGYTPLIVV 240
                *.::*:**:***::*:**.****::******:.:.**:*::****.****:*.*:**::*

YiaO_Ecoli      INKAKFDGLSPEFQQALVSSAQEAGNYQRKLVAEDQQKIIDGMKEAGVEVITDLDRKAFS 299
YiaO_Styph      INKAKFDALSPQFQEALLSSAKEAGDYQRKLVAEDQQKIIDGMKEAEVEVLTDIDRKAFS 299
YiaO_Hinfl      MNKAKFDSLLPALQTAIIEAAKEAGQFQRDLNVKNEQNIISKLRKQGVEVIEKINTEPFK 300
PM1252_Pmult    MNKAKFDGLSPALQSAILEAAKEAGAYQRKLNLDNEKGIIEKMQKAGIQVIETVDTKPFK 300
                :******.* * :* *::.:*:*** :**.*  .::: **. :::  ::*:  :: :.*.

YiaO_Ecoli      DALGNQVRDMFVKDVPQGADLLKAVDEVQ- 328
YiaO_Styph      DALGSQVRDMFLKDNPQGADLLKAVDEVQ- 328
YiaO_Hinfl      TLIEEKVRKSFIEK--HGDDLLKKVDALSE 328
PM1252_Pmult    AAIESEVRKAFIEK--NGDDLVKQIDALAK 328
                  : .:**. *::.  :* **:* :* :
```

**Figure 4.** Alignment of *E. coli* K-12 YiaO (YiaO_Ecoli) with three close homologs from *S. typhimurium* (YiaO_Styph), *H. influenzae* (YiaO_Hinfl) and *P. multocida* (PM1252_Pmult). The N-terminal amino acid sequence that was determined for the purified *E. coli* YiaO is underlined. The predicted signal sequence for targeting and processing in the periplasm is indicated in bold face for all proteins.

## Discussion

To perform a detailed study of transport protein mechanism, several tools are required. Ideally, the transporter is overexpressed and purified in sufficient amounts to allow active reconstitution into liposomes. This way the kinetics of transport can be studied without the interference of metabolic activities of the cell. In addition, the requirements for putative accessory proteins and the nature of the driving force can be studied. Various examples exist of successful efforts with transporters that consist of a single [103, 133, 168] or two [209, 212] protein subunits. The elaborate nature of these studies shows that successful conditions have to be carefully determined for each transporter individually.

Only a few reports have appeared on the reconstitution of multi-component ABC transporters, consisting of three or more distinct subunits. The best studied examples to date are the maltose transporter $MalFGK_2$/MBP of *E. coli* [47], and the histidine permease $HisJQMP_2$ of *S. typhimurium* [118]. Recently, the purified components of the *E. coli* vitamine $B_{12}$ transporter Btu have been used to model the interaction of the substrate-binding protein BtuF with its cognate transporter domains, albeit in detergent solution and not in a reconstituted membrane system [27]. Attempts have been made to obtain an active, purified and reconstituted complex of both the GlcTUV/GBP glucose transporter of the Archaeon *Sulfolobus solfataricus* [4], and the OppABCDF oligopeptide transporter of the Gram-positive bacterium *L. lactis* [49], but these attempts were unsuccessful. In the latter case, the membrane and ATPase subunits could not be overexpressed to sufficient amounts, either separately or in varying combinations. The experiments did show, however, that co-expression of the subunits significantly increased protein yield [49]. A similar phenomenon has been reported for expression of the essential membrane component SecY of the bacterial translocase, which is only stable when SecE is co-expressed [100, 138]. In the absence of SecE, SecY is degraded by the membrane-bound protease FtsH.

Strikingly, we observed a similar phenomenon when overexpressing the membrane domains of the YiaMNO transporter. Low-level expression of His-YiaM could only be detected if at least both the *His-yiaM* and *yiaN* genes were expressed (Fig. 2).

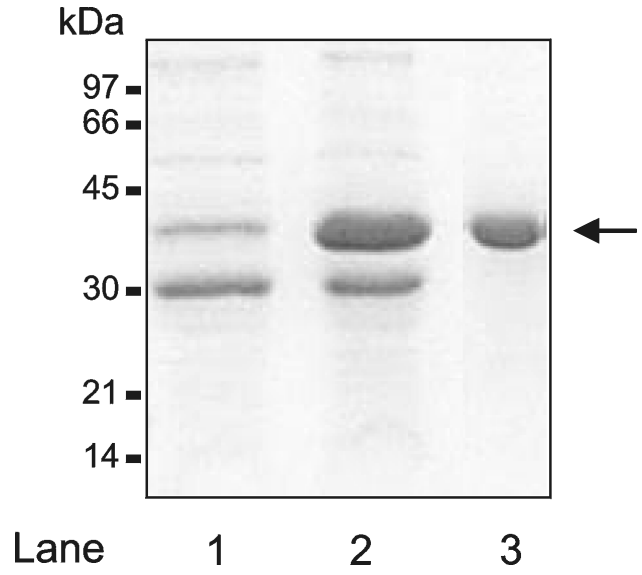



**Figure 5.** Overexpression and purification of His-tagged YiaO. C-terminally 6xHis-tagged YiaO (YiaO-His, arrow) was overexpressed in the periplasm of *E. coli* SF100 (lane 1: pSA5 (empty vector), lane 2: pET917 (expression vector)). YiaO-His was purified from the periplasmic fraction using Ni2+-NTA affinity chromatography (lane 3). Proteins were analyzed by 15% SDS-PAGE followed by CBB-staining.

The procedure has not been optimized and no further efforts were made to purify the transporter since no means for assaying the activity was available (see below). The substrate-binding protein YiaO could readily be overexpressed and purified from the periplasmic fraction of overexpressing cells, using size-exclusion chromatography (Fig. 3). N-terminal signal sequence determination of the 36-kDa protein not only confirmed its identity, but also revealed that the protein is processed at the signal sequence cleavage site that is predicted for Gram-negative bacteria (Fig. 4) [149]. Purification of YiaO-His via a C-terminal His-tag was not only more rapid but also more efficient, since the 30-kDa major contaminant of the size-exclusion procedure (Fig. 3) does not co-purify during $Ni^{2+}$-NTA affinity chromatography (Fig. 5) [Chapter 4]. Therefore, in future studies requiring purified YiaO, the latter procedure was followed.

To monitor transport activity, it is essential that the substrate is available in a radiolabeled or fluorescent form. Unfortunately, L-xylulose, the only known substrate for the YiaMNO transporter [Chapter 4], cannot be obtained in radiolabeled form nor is any labeled competitor available that would facilitate indirect binding and/or transport studies. Although an enzyme assay was developed that can be used to monitor the L-

xylulose content of supernatants, this method is not sensitive enough for use in uptake assays with membrane vesicles or proteoliposomes. Since the activity of the YiaMNO transporter could not be monitored during the steps of the solubilization, purification, and reconstitution procedure, reconstitution of the YiaMNO transporter was not attempted.

Both the function and transport mechanism of the binding protein-dependent secondary transporters known to date still pose intriguing questions. Up to now, the tools that are essential for detailed investigation of the properties of these transporters are lacking, which greatly hinders our attempts to learn more about the function of these proteins. Nevertheless, as the family expands with every completely sequenced prokaryotic genome that contains one, as is the case for *E. coli* K-12, up to a dozen, in *Bordetella* [10], of these transporters, more research will be devoted to the function of its members.

# *Chapter 3*

## On the physiological role of the *yiaMNO* transporter genes of *Escherichia coli* K-12

*"my feeling is that there are dark forces at work here"* [245]

Titia H. Plantinga, Chris van der Does, Danuta Tomkiewicz, Geertje van Keulen[1], Wil N. Konings and Arnold J.M. Driessen

**Summary**

Binding protein-dependent secondary, or tri-partite ATP-independent periplasmic (TRAP), transport proteins make up a unique transport protein family, since the binding protein-dependent uptake of solutes is driven by the proton-motive force rather than ATP-hydrolysis. This family was described only recently and few of the systems have been investigated regarding their physiological function. The *yiaMNO* genes of *Escherichia coli* K-12 encode a transporter belonging to this family that was implicated in the uptake of the rare pentose L-xylulose. However, the physiological role of this transporter is unknown as wild type *E. coli* K-12 strains do not utilize L-xylulose as the sole carbon source, nor is expression of the transporter induced by this carbohydrate. The *yiaMNO* genes are, however, expressed in a growth phase-dependent manner in *E. coli* K-12 strain MC4100. Deletion of the genes in this strain resulted in remarkable changes in its growth characteristics, including the transition from exponential growth to the stationary phase, high salt-survival and biofilm formation.

[1]*Microbial Physiology Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands*

## Introduction

Prokaryotes employ several classes of transport systems for the uptake of solutes from their environment, which are defined on the basis of their subunit composition and mode of energization [54]. A recent addition to the list of transport protein classes is the family of binding protein-dependent secondary [53, 54], or tripartite ATP-independent periplasmic (TRAP) [59, 98, 171], transporters. These systems involve a solute-binding protein that captures the substrate at the outside of the cell, and subsequently delivers it to an integral membrane permease that is made up of two dissimilar subunits. The large subunit contains 12 putative transmembrane domains (TMDs) and a large cytoplasmic loop between TMD 6 and TMD 7, and thus resembles the classical secondary transporters. The small subunit is made up of 4 putative TMDs. Transport across the cytoplasmic membrane is driven by the proton-motive force (pmf) [53, 54, 59, 98, 171, 241].

To date, only a few members of this transport protein family have been investigated regarding their physiological function. The first to be described both in molecular and biochemical detail was the DctPQM transporter of *Rhodobacter capsulatis* [59]. This system is involved in the uptake of the $C_4$-dicarboxylates malate, succinate, and fumarate [59]. A similar system was found in *Wolinella succinogenes* [210]. Biochemical evidence indicates that one member is involved in the sodium- and pmf-dependent uptake of glutamate by *Rhodobacter sphaeroides* [91]. More recently, TeaABC of *Halomonas elongata* was shown to transport the compatible solutes ectoine and hydroxyectoine [66]. The latter finding clearly demonstrates the involvement of this type of transporter both in the uptake of carbon sources and the protection of the cell against unfavorable conditions.

The *Escherichia coli* K-12 genome contains one member of this class of transporters. The *yiaMNO* genes, that are located in the *yiaKLMNOPQRS* gene cluster, encode a transporter that mediates the uptake of the rare pentose L-xylulose (L-*threo*-2-pentulose) and possibly other sugars [Chapter 4]. YiaO specifies the periplasmic binding protein, while YiaM and YiaN constitute the integral membrane permease. L-xylulose is presumably metabolized by the enzymes encoded by genes located immediately downstream of the *yiaMNO* genes, since L-xylulose transport and metabolism is found

only in cells that constitutively express the *yiaK-S* gene cluster. L-xylulose, however, does not induce expression of these genes, and little is known about the natural occurrence and abundance of the pentose and it's utilization by *E. coli* K-12 [87, Chapter 4]. Therefore, L-xylulose might not be the sole substrate for the YiaMNO transporter.

This study focuses on the physiological function of the *yiaMNO* genes. In previous studies, the *yiaK-S* cluster was found to be repressed in wild type *E. coli* strain ECL1 [87, Chapter 4]. Therefore, expression of the *yiaMNO* genes was analyzed in *E. coli* strain MC4100, which is widely used in studies on the *E. coli* K-12 physiology. An unmarked chromosomal deletion of the *yiaMNO* genes was constructed in strain MC4100, which resulted in some marked phenotypical effects. The transition from exponential to stationary growth was delayed, resulting in higher final cell densities. In addition, high-salt survival was impaired, and biofilm formation reduced.

## Experimental procedures

*Bacterial strains and growth conditions*

All used strains are *E. coli* K-12 derivatives: MC4100 ($F^-$ *araD*139 $\Delta$(*argF-lac*)*U*169 *rpsL*150 *relA*1 *flbB*5301 *deoC*1 *ptsF*25 *rbsR*; laboratory collection), JA134 (*hfrC phoA*8 *relA*1 *tonA*23 $T2^r$ *lyx*$^+$) [183], and MG1655 ($F^-$ $\lambda^-$ *ilvG rfb*50 *rph*1; laboratory collection). The unmarked chromosomal deletion of the *yiaMNO* genes in these strains was created as described elsewhere [Chapter 4], yielding mutant strains TP001, TP018 and TP007, respectively. Strains were grown aerobically in Luria-Bertani (LB), LB supplemented with 0.5% (w/v) glucose (LBG), or M63 minimal medium [15, 141] at 37°C. Growth was monitored by optical density measurements at 660 nm ($OD_{660}$). Colony forming units (cfu's) were determined by plating serial dilutions on LB-agar. Antibiotics were used at: ampicillin: 50 μg/ml, tetracyclin: 12 μg/ml. For high-salt growth-experiments, LB and LBG were supplemented with 0.8-1.0 M NaCl or KCl.

*Expression of the* yiaM *and* yiaO *genes*

Cells were grown aerobically in LB and LBG, samples were taken at various $OD_{660}$'s and total RNA was extracted as described [Chapter 4]. RT-PCR was performed on 1 μg total RNA per experiment, using RT-PCR beads (Amersham Pharmacia Biotech) and primers directed against 474- and 632-bp fragments of the *yiaM* and *secY* genes, respectively [Chapter 4]. For Northern Blot analysis, 5 μg total RNA from each sample was run on a 1.2% (w/v) agarose gel containing 6.7% (v/v) formaldehyde and transferred to a positively charged nylon membrane (Zeta Probe, Biorad) in 150 mM NaCl, 15 mM $Na_3$-citrate. Transfer

was checked by ethidium bromide staining of the membrane. For preparation of the probe, the *yiaO* gene was cloned via PCR, using forward primer 5'-AATGGATCCATTAAAAGGAAAATATTATG-3' and reverse primer 5'-CCCTCTAGATTATTGCACCTCATCCAC-3', introducing *Bam*HI and *Xba*I restriction sites, respectively. The fragment was ligated into vector pET401 [214] for propagation, which was used as a template for labelling with $^{32}$P-dCTP with Klenow polymerase (Roche). The probe was purified using a PCR-product isolation procedure (Qiagen), and hybridized overnight in 0.5 M sodium phosphate pH 7.2, 1% (w/v) blocking reagent (Roche) and 7% (w/v) SDS at 65°C. The membrane was washed in 50 mM sodium phosphate pH 7.2 containing 1% SDS, and the signal was recorded by autoradiography.

*Promoter induction and β-galactosidase assays*

The 1000-bp upstream region of *yiaM* was cloned via PCR. Forward primer 5'-ATGGTGGATCCGATGATGAGGGCA-3' introduced a *Bam*HI-site, and reverse primer 5'-TGAATTCATAGCTATTCCTTGAGGC-3' introduced an *Eco*RI-site. The fragment was translationally fused to the *lacZ*-reporter gene in vector pBC3 [139]. This vector was labeled pP1000 and transformed into strain MC4100. Heat and cold shock were applied by shifting liquid cultures to 42°C and 10°C, respectively. High salt, sucrose, and spent medium (prepared as described below) effects were determined by harvesting the cells and resuspending them in the respective media. Following 30 minutes of incubation, the cells were harvested, and β-galactosidase assays were performed as described [142]. Values were expressed in Miller Units (MU) [142].

*Spent medium growth experiments*

MC4100 and TP001 cells were grown aerobically in LB or LBG, 50 ml samples were taken at $OD_{660}$ values 0.5, 1.5, 2.0 and 2.5, and harvested (4,000 x g, 10 min, 4°C). The supernatant was filtered (FP30/0,2 0.2 μm pore-size sterile filters, Schleicher and Schuell) to remove all cells. Freshly inoculated LB or LBG cultures of MC4100 and TP001 were grown to an $OD_{660}$ of 0.5, harvested, and resuspended in equal volumes of pre-warmed (37°C) spent medium. Aerobic growth was continued and monitored over time.

*Nuclear magnetic resonance analysis*

MC4100 and TP001 were grown to an $OD_{660}$ of 0.5 in a total of 8l LBG containing 0.9 M NaCl per strain. Cells were harvested and freeze-dried overnight. Ethanol extracts for determination of intracellular solutes were prepared as described elsewhere [135]. $^{13}$C-nuclear magnetic resonance (NMR) spectra were measured by H. Santos at the Instituto Technologia Química e Biológica of the University of Lisbon (Portugal) as described [135] using a Bruker DRX500 spectrometer.

*Auto-aggregation and biofilm experiments*

Auto-aggregation of the strains was assessed as described [101] with modifications: overnight cultures were adjusted to identical $OD_{660}$ values, divided into 10 ml-samples in sterile 14-ml tubes, and left standing at room temperature. Samples were taken from the top 0.5 cm of the tubes and the $OD_{660}$ was measured. For biofilm experiments, cells were plated on LB-agar and grown overnight at 37°C. Colonies were picked and resuspended in M63 minimal medium, and diluted into M63-filled 96-wells untreated polystyrene plates (Costar) containing 0.5% (w/v) glucose. Plates were incubated at 37°C for 60h, and biofilm formation was quantified using crystal violet, as described elsewhere [151].

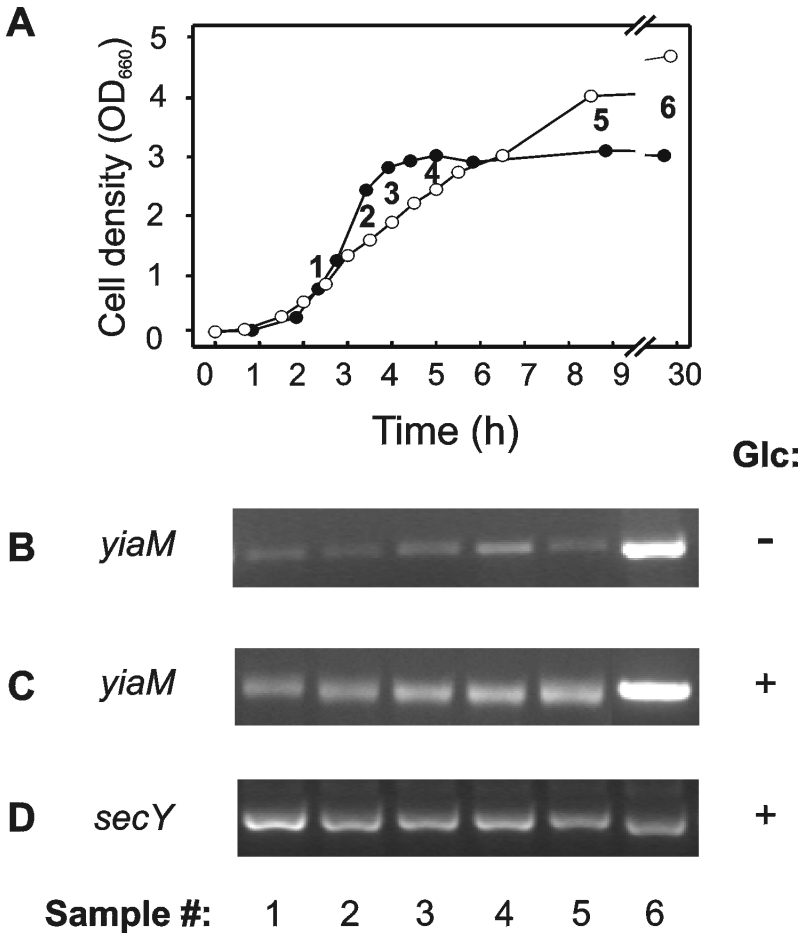



**Figure 1.** Expression of the *yiaM* gene at different growth stages. (A) MC4100 cells were grown aerobically in LB (open circles) and LBG (closed circles). Total RNA was isolated at indicated growth phases: 1: exponential; 2: early, 3: mid, and 4: late transition, 5: early stationary phase, 6: 24 hrs stationary. (B) Expression pattern of the *yiaM* gene in LB, and (C) in LBG. (D) *secY*: control for constitutive expression.

## Results

*Expression of the system in* E. coli *strain MC4100*

The physiological role of the *yiaMNO* transport protein genes was mainly studied in strain MC4100. This strain expressed the *yiaM* gene during growth both in LB and LBG, as demonstrated using RT-PCR (Fig. 1B, C). Expression appeared to be growth phase-dependent, and was maximal once the cells had been in stationary phase for several hours (Fig. 1B, C). Expression was not repressed by glucose (Fig. 1C). These results were supported by Northern Blotting, using the *yiaO* gene as the hybridization probe (data not shown).

*E. coli* strain JA134 expresses the *yiaK-S* gene cluster from one promoter upstream of the *yiaK* gene [87]. Computational analysis of this region on the *E. coli* K-12 genome, using RegulonDB [182] identified an additional putative promoter immediately upstream of the *yiaM* gene (Table 1). This putative promoter was translationally fused to a *lacZ*-reporter construct and it's activity in strain MC4100 was determined. The low-level expression pattern confirmed the RT-PCR and Northern Blotting results (data not shown), indicating that the *yiaM-S* promoter was active. However, the promoter was not induced by L-xylulose (data not shown), as is the case for the *yiaK*-promoter [87].

**Table 1.** Nucleotide sequence of the putative *yiaM-S* promoter

| Promoter | -35 region | (# bp) | -10 region |
|---|---|---|---|
| *yiaM-S* | C G A C T T G A A T | 12 | C C T C A A G G A A T |
| σ 32[a] | - C C C T T G A A - | 13-15 | - C C C C A T - T A - |
| σ 70[b] | - - T C T T G A C A | 16-18 | - - - - A T A T A A T |

Black and grey boxes indicate identical bases are found on these positions.

a taken from ref. 172

b taken from ref. 246

*Deletion of the* yiaMNO *genes affects growth*

A *yiaMNO* deletion mutant of MC4100 was constructed and labeled TP001. Expression of the genes located immediately up- and downstream of the transporter was not affected (data not shown), as has been shown for TP018, the Δ*yiaMNO* derivative of JA134 [Chapter 4]. MC4100 and TP001 have been used in an extensive search for the substrate of the YiaMNO transporter (see Chapter 4). Although these experiments did not identify the transported substrate, a phenotypical difference between mutant and parent strain was observed. When grown in LBG, both strains grew at a nearly identical rate during the exponential phase, with the relatively sharp transition from exponential to stationary phase which is typical for a carbon-limited batch-culture [137]. However, for TP001 this transition was delayed, reproducibly yielding a higher final $OD_{660}$ at stationary phase (Fig. 2). This difference in growth was also observed when the cells were growing in minimal media supplied with carbon sources other than glucose (data not shown, compounds listed in Chapter 4), but was absent in LB without glucose (Fig. 2).

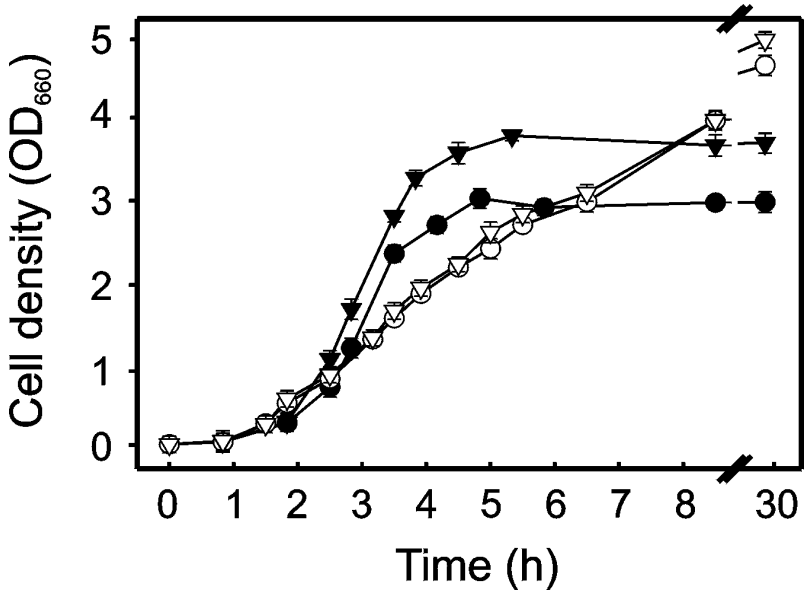



**Figure 2.** Effect of the *yiaMNO*-deletion on growth in rich media. When grown on LBG (filled symbols), strain TP001 (triangles) reproducibly reached a higher final cell density then MC4100 (circles), whereas there were no differences in growth on LB (open symbols).

The difference in growth was supported by cfu-numbers (data not shown). These findings suggest that, under specific conditions, the deletion of the three structural genes encoding the YiaMNO transporter confers a growth advantage on these cells.

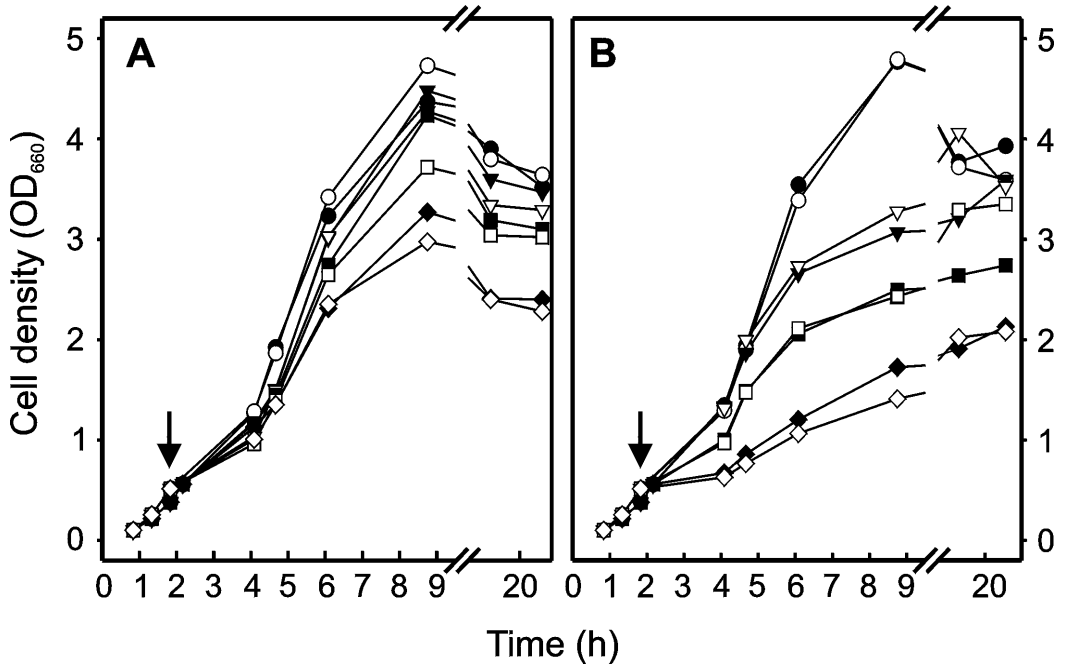



**Figure 3.** Effect of MC4100-spent medium on growth. Fresh cultures of MC4100 (closed symbols) or TP001 (open symbols) were grown in (A) LB or (B) LBG to an $OD_{660}$ of 0.5 (arrow), harvested, and resuspended in cell-free culture supernatants. Spent medium was prepared from cultures grown on either LB or LBG to $OD_{660}$'s of 0.5 (circles), 1.5 (triangles), 2.0 (squares) and 2.5 (diamonds).

*Involvement of the YiaMNO transporter in quorum sensing*

Interestingly, production of autoinducer 2 (AI-2), the quorum sensing signal molecule of *E. coli* and a range of other bacteria [37], occurs at the conditions where the growth advantage of the Δ*yiaMNO* strain was observed [198]. Therefore, we investigated whether the transporter, or the transported substrate(s), could be involved in quorum sensing by examining the effect of spent medium on growth. Spent medium taken from cultures grown in LB hardly influenced the growth of fresh cultures (Fig. 3A). In contrast, addition of LBG spent medium taken at an $OD_{660}$ > 1.5 clearly negatively affected growth (Fig. 3B). However, no differences in response were observed between MC4100 and TP001 (Fig. 3). Spent medium prepared from TP001 gave identical results

(data not shown). In addition, no induction of the *yiaM-S* promoter by spent medium was observed (data not shown). These findings indicate that if the autoinducer indeed is produced by these strains during growth on LBG, the ability of TP001 to respond to the molecule has not been affected by the deletion of the *yiaMNO* genes.

*Deletion of the* yiaMNO *genes affects high-salt tolerance*

The *H. elongata* TeaABC transporter protects the cell against hyperosmotic conditions [66]. Therefore, the ability of strains MC4100 and TP001 to survive hyperosmotic stress was investigated. In the presence of 0.9 M NaCl, which is 10-times the concentration of NaCl in LB(G), the deletion strain showed a marked lag time (1 h 23 min ± 12 min, n=6) compared to the parent strain before resuming growth. The strain did not grow at all at 1 M NaCl (Fig. 4). Similar results were found when the cells were grown in the presence of high KCl concentrations, but not in media containing sucrose (up to 30%, data not shown). None of these conditions induced the *yiaM-S* promoter (data not shown).

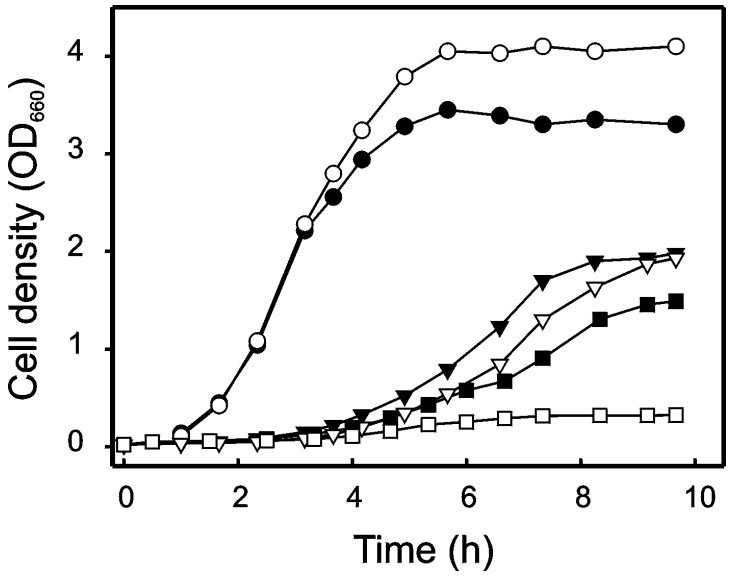



**Figure 4.** High salt sensitivity of the *yiaMNO*-deletion mutant TP001. Both parent MC4100 (filled symbols) and mutant TP001 (open symbols) were grown in LBG containing 100 mM (normal concentration; circles), 0.9 M (triangles) or 1 M (squares) NaCl. The ability of TP001 to respond to high salt stress decreased with increasing salt concentration.

$^{13}$C-NMR analysis of whole-cell extracts obtained from both *E. coli* strains grown in LBG containing 0.9 M NaCl detected no differences in total accumulated cellular compounds. The major compatible solute that had been accumulated by both strains was identified as glycine betaine (data not shown). In this regard, none of the compatible solutes tested in our previous experiments, i.e. glycine betaine, ectoine, and K-glutamate, were found to be substrates for the YiaMNO transporter [Chapter 4].

*Deletion of* yiaMNO *reduces biofilm formation*

We investigated whether additional properties had been affected by the *yiaMNO* deletion. Indeed, autoaggregation of TP001 cells in static liquid culture was dramatically reduced compared to that of MC4100 (Fig. 5). This prompted us to investigate biofilm formation by this strain, since both bacterial aggregation and adhesion play important roles in this process [115, 152, 155]. Biofilms were allowed to form for 60 h in M63 minimal medium in the presence of D-glucose as the sole carbon source. As observed before, TP001 reached a higher final $OD_{660}$ compared to MC4100 during growth (data not shown). However, biofilm formation was negatively affected by deletion of the YiaMNO transporter, since significantly less TP001 biomass was attached to the surface of the wells compared to parent strain MC4100 (Fig. 6).

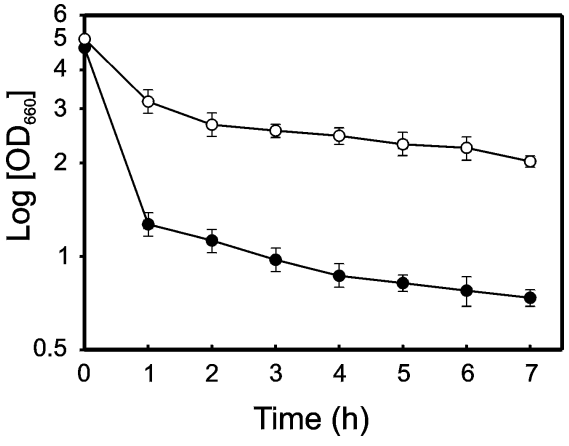



**Figure 5.** Autoaggregation of the cells in static cultures was investigated by monitoring the $OD_{660}$ of the top layer of static liquid cultures in time. Sedimentation of deletion mutant TP001 (open circles) is dramatically reduced compared to that of MC4100 (closed circles).

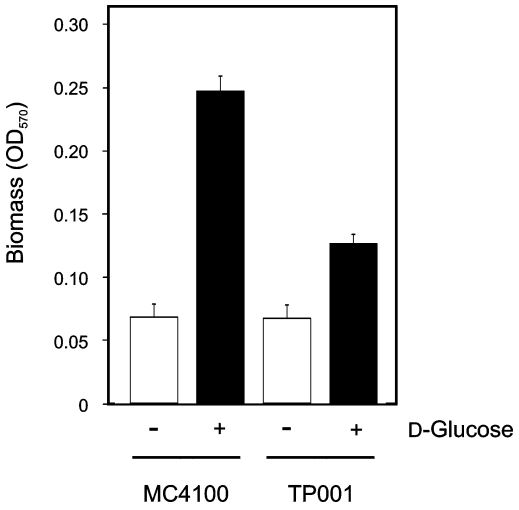



**Figure 6.** Biofilm formation detected by crystal violet staining of attached biomass. Cells were grown for 60h in M63 minimal medium in the absence (-) or presence (+) of D-glucose. In the presence of D-glucose, both strains form biofilms, but biofilm production is reduced in the deletion mutant TP001 compared to the parent strain MC4100.

## Discussion

This study addresses the physiological function of the *yiaMNO* genes, making use of strain MC4100 that expresses the *yiaM* and *yiaO* genes in a growth phase-dependent manner (Fig. 1). Expression of the *yiaK-S* gene cluster in strain JA134 is regulated by cyclic AMP receptor protein (CRP), and thereby subject to carbon catabolite repression [87]. The *yiaM-S* promoter that was identified by computational analysis of the *E. coli* K-12 genome contains a putative CRP-binding site [203], but the system was not repressed by glucose in strain MC4100 (Fig. 1). The promoter was not induced by any of the applied physiological conditions, and no *yiaK-S* expression has been reported in the available *E. coli* expression profile data [11, 12, 162, 204, 233]. Thus, these expression data provide no functional information about the system. Nevertheless, highest expression levels were detected in cells that had been in stationary phase for 24h (Fig. 1). As the stationary growth phase under laboratory conditions more closely resembles the natural situation for bacteria than the exponential growth phase [105], the YiaMNO transporter may play a role in scavenging scarce substrate under limiting conditions.

The loss of the YiaMNO transporter resulted in an apparent growth advantage on LBG (Fig. 2). This effect of the deletion was observed under conditions where the

quorum sensing signal AI-2 is produced [19, 37, 198], and deletion of the AI-2 producing enzyme results in faster cell growth [195]. Although our data provide the first indication that strain MC4100 produces an 'AI-2-like' activity, we find no evidence for involvement of YiaMNO in transport of the autoinducer AI-2 or a precursor thereof (Fig. 3). The deletion proved to be a disadvantage at other conditions, as high-salt survival (Fig. 4), autoaggregation, an initial step of biofilm formation [115, 152, 155] (Fig. 5), and biofilm formation itself by TP001 were significantly reduced compared to MC4100 (Fig. 6).

Strikingly, an increase in fitness that may be related to deletion of the *yiaMNO* genes has been reported for the Gram-negative bacterium *Ralstonia* sp. TFD14. After 1000 generations of experimental evolution, all evolved populations showed an increased fitness compared to the ancestor strain [108], and 71 out of 72 evolved populations had lost the 2,4-kb genomic fragment containing the *Ralstonia yiaMNO*-homologues [146]. Since the size of the complete deletion was not determined, the possibility cannot be excluded that one or more up to now unidentified gene(s) contribute(s) to the phenotype in this organism. In addition, the elaborate *Ralstonia* EPS had disappeared, leading to changes in bile salt-sensitivity and adhesion behavior [176]. The major EPS of *E. coli* K-12, colanic acid [175], is required for development of biofilm architecture, and production is increased during this process [39, 167]. Deletion of the producing genes delays, but does not abolish, biofilm formation, with nearly identical amounts of surface-attached biomass at times ≥ 45h [39]. This indicates that the *yiaMNO* deletion most likely does not interfere with colanic acid biosynthesis, since the differences in biofilm formation between parent and deletion mutant were still observed after 60h (Fig. 6). Nevertheless, although the transported substrate L-xylulose is described neither as a constituent, nor as a precursor for *E. coli* K-12 EPS biosynthesis [175, 197], this possibility cannot be excluded at this time.

Regulation of biofilm development by *E. coli* K-12 has been studied to some detail. Most studies make use of *ompR* mutants of strain MC4100, since, in contrast to our experiments, in these studies the wild type strain does not form a biofilm [166, 233]. The *ompR* mutation increases expression of curli, a fibrillar structure on the cell-surface that increases adherence properties of the cell [233]. Other studies link the envelope stress Cpx signal transduction pathway [52, 155] to the process. Although in our hands

MC4100 forms a biofilm in the presence of glucose (Fig. 6), other studies report it is subject to catabolite repression [89, 90].

Δ*yiaMNO* derivatives of two additional *E. coli* K-12 strains were used to check whether the observed phenotype was specific for this deletion. The effects of the deletion were not observed with *yiaMNO*-deletion mutants of strain JA134, that was used to identify the substrate [Chapter 4], and the completely sequenced wild type strain MG1655 [22] (T.H. Plantinga, C. van der Does, W.N. Konings, and A.J.M. Driessen, unpublished results). However, both strains differed in their growth and aggregation behavior from strain MC4100 and therefore the effect of the deletion may have been obscured. Nevertheless, biofilm formation by strain MG1655 was negatively affected by the deletion (T.H. Plantinga, C. van der Does, W.N. Konings, and A.J.M. Driessen, unpublished results). No *yiaMNO*-expression was detected in MG1655 biofilms, but it must be noted that these experiments were conducted under circumstances that differed from our experiments [187].

Since L-xylulose is unable to induce the expression of the YiaMNO transporter, while expression is not repressed by glucose, a major role of this system in carbon source uptake and utilization is not evident. Nevertheless, the system may play a role in scavenging of scarce substrate during stationary phase. The observed phenotypic effects are rather diverse, and the pathways that underlie these phenomena are complex. However, the observations made with the *Ralstonia* sp. clones that have lost the *yiaMNO* genes support our finding that a localized deletion may have drastic effects on whole cell physiology. Further studies are required to establish the physiological role of the binding protein-dependent secondary transporter encoded by the *E. coli* K-12 *yiaMNO* genes.

## Acknowledgements

We thank Henk Bolhuis and Evelien te Poele (Microbial Ecology group, University of Groningen) for helpful discussions, and Ana Ramos and Helena Santos (Instituto Technologia Química e Biológica, University of Lisbon) for the NMR-data. This work was supported by the Netherlands Organization for Scientific Research (NWO, grant no. 805-19-046 P).

# *Chapter 4*

## Functional characterization of the *Escherichia coli* K-12 *yiaMNO* transport protein genes

*"When you have eliminated the impossible, whatever remains, however improbable, must be the truth."* [38]

Titia H. Plantinga, Chris van der Does, Josefa Badia[1], Juan Aguilar[1], Wil N. Konings and Arnold J.M. Driessen

(This chapter has been accepted for publication in *Molecular Membrane Biology*)

### Summary

The *Escherichia coli* K-12 genome contains one member of the recently discovered family of binding protein-dependent secondary, or tri-partite ATP-independent periplasmic (TRAP), transporters. Since only a few members of this family have been functionally characterized to date, we aimed to identify the substrate of the *E. coli* homologue encoded by the *yiaMNO* genes. More than 100 potential substrates for the YiaMNO transporter were tested. We found that cells constitutively expressing the *yiaK-S* gene cluster metabolize the rare pentose L-xylulose whereas the parent strain does not. Deletion of the *yiaMNO* transporter genes reduced L-xylulose metabolism. Since L-xylulose cannot be radioactively labeled, indirect methods were used to measure substrate binding and transport. Thus, we found that the periplasmic substrate-binding protein YiaO binds L-xylulose, and stimulates consumption of the pentose by spheroplasts that express the membrane domains YiaM and YiaN.

[1] *Department of Biochemistry, School of Pharmacy, University of Barcelona, Barcelona, Spain*

## Introduction

Prokaryotes use a large variety of membrane-localised transport proteins to facilitate uptake of solutes from their environment, which are divided into distinct classes based on energy requirement and polypeptide composition [54]. Binding protein-dependent secondary transporters [53, 54, 91], or tri-partite ATP-independent periplasmic (TRAP) transporters [59, 98, 171] form a new class of transporters. These systems share characteristics with both ATP-binding cassette (ABC) [8, 77, 79] and secondary [129, 163] transporters. Uptake of solutes involves an extracytoplasmic solute-binding protein, but the driving force is provided by the proton- (pmf) and/or sodium ion motive force (smf) rather than ATP hydrolysis. Typically, the membrane domain consists of two dissimilar proteins. The large subunit consists of 12 putative transmembrane domains (TMDs) and a large central cytosolic loop and resembles classical secondary transporters, and the smaller subunit consists of 4 putative TMDs [53, 171, 241]. The architecture and biochemical characteristics of these carriers pose intriguing questions regarding both the mechanistic properties and the evolution of transport proteins [54]. To date, only a limited number of these proteins have been described in molecular detail.

Biochemical evidence for the existence of these systems was provided by studies on glutamate uptake in *Rhodobacter sphaeroides* [91]. The transport protein family was defined when the *dctPQM* genes of *Rhodobacter capsulatus* were shown to encode a binding protein-dependent secondary uptake system for the $C_4$-dicarboxylates malate, succinate, and fumarate [59]. Additionally, the *dctPQM* homologues of *Wolinella succinogenes* were linked to $C_4$-dicarboxylate uptake [210]. Homologous systems are found in all bacterial subdivisions as well as in archaea [53, 59, 98, 171]. Recently, a new family member was discovered that transports the compatible solutes ectoine and hydroxyectoine in *Halomonas elongata* [66].

The *Escherichia coli* K-12 *yiaMNO* genes encode a binding protein-dependent secondary transporter [22, 53, 59, 98, 171], see also the alignments presented on: http://www-biology.ucsd.edu/~msaier/transport/phylo/trap.html. The genes are located within the *yiaK-S* cluster (GenBank accession nr. g1789999-08, Fig. 1A), that has been implicated in carbohydrate utilization [17, 87]. Strain JA134 constitutively expresses

these genes due to a genomic rearrangement, which enables these cells to grow on the pentose L-lyxose [17, 183]. The gene located immediately downstream of *yiaMNO* encodes LyxK, a kinase that phosphorylates the rare pentose L-xylulose (L-*threo*-2-pentulose) [16, 183]. This suggests that the YiaMNO transporter may be involved in carbohydrate uptake.

To provide more insight into the function of the YiaMNO transporter, we have carried out a deletion and biochemical analysis of the system in strain JA134. These cells are able to metabolise L-xylulose, whereas in *yiaMNO* deletion mutants this ability is reduced. In addition, we find that YiaO binds L-xylulose and together with the YiaM and YiaN proteins can function as a binding protein-dependent uptake system for this pentose.

**Table 1.** *E. coli* K-12 strains used in this study

| Strain | Characteristics | Source or reference |
|---|---|---|
| MC4100 | *araD139 Δ(argF-lac)U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR* | Laboratory collection |
| TP001 | MC4100 Δ*yiaMNO* | This work |
| ECL1 | HfrC *phoA8 relA1 tonA22* $T2^r$ (λ) | 116 |
| JA134 | ECL1 $lyx^+$ | 183 |
| TP018 | JA134 Δ*yiaMNO* | This work |
| EC1000 | MC1000 $repA^+$ | 114 |
| SF100 | KS272 Δ*ompT* | 18 |
| XL1-Blue | *recA*1 *lac hsdR17 supE44 relA1* (F' *proAB* $lacI^q$ *lacZ* dM15 Tn*10*) | Strategene |
| TE2680 | $F^-$ $\lambda^-$ IN (*rrnD-rrnE*) Δ(*lac*)*X74 rplS galK2 recD::*Tn*10d-tet trpDC700::putPA1303*:: [$Km^s$ $Cm^r$ *lac*] | 56 |

## Experimental procedures

*Bacterial strains, plasmids, primers and growth conditions*

*E. coli* K-12 strains, vectors and recombinant plasmids used in this study are listed in tables 1 and 2, respectively. PCR and RT-PCR primers are listed in table 3. Physiological studies were performed with *E. coli* MC4100, TP001, ECL1, JA134, and TP018. DH5α was used for cloning. EC1000 was used for handling pORI240 and its derivatives. SF100 was used for overexpression of His-tagged YiaO. The strains were grown aerobically at 37°C in Luria Broth (LB) or in M63 minimal medium [15]. Antibiotics were added to final concentrations: ampicillin (Am) 50 μg/ml; tetracyclin (Tc) 12 μg/ml; kanamycin (Km) 50 μg/ml. 5-Bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) was used at 30 μg/ml.

**Table 2.** Plasmids used in this study

| Plasmid | Characteristics | Reference |
|---|---|---|
| pET401 | cloning vector, $Am^r$, *trc*-promoter | 214 |
| pET908 | pET401 carrying fused F1-F2 flanking regions | This work |
| pORI240 | $Tc^r$, $LacZ^+$, $ori^+$ of pWV01, requires *repA* *in trans* for replication | 114 |
| pORIF1F2 | pORI240 carrying fused *yiaMNO*-flanking regions | This work |
| pRS550 | promoter-less *lac*, $Km^r$, $Am^r$ | 191 |
| pSA5 | Expression vector, $Am^r$, C-terminal 6xHis-tag | 6 |
| pET917 | pSA5 carrying *yiaO* | This work |

*Search for the substrate of YiaMNO*

The following compounds were tested as possible substrates of the transporter in growth experiments, metabolic assays (tetrazolium violet and/or Biolog Microplate assays described below) and, when available in radioactive form, in uptake and binding experiments: (miscellaneous) Tween-20, Tween-40, Tween-80; ($C_{18}$) maltotriose; ($C_{12}$) cellobiose, α-D-lactose, lactulose, maltose, D-melibiose, sucrose, D-trehalose; ($C_{10}$) adenosine, 2'-deoxy adenosine, inosine, thymidine; ($C_9$) uridine; ($C_8$) N-acetyl-D-glucosamine, N-acetyl-β-D-mannosamine, *m*-hydroxy phenylacetic acid, *p*-hydroxy phenylacetic acid, phenylethylamine, tyramine; ($C_7$) glycyl-L-glutamic acid, glycyl-L-proline, α-methyl galactoside, β-methyl glucoside; ($C_6$) L-ascorbate, $Fe_3$-citrate, iso-citrate, 2,3-diketo-L-gulonate, dulcitol, ectoine, D-fructose, fructose-6-phosphate, L-fucose, D-galactonic acid γ-lactone, L-galactonic acid γ-lactone, D-galactose, D-galacturonic acid, D-gluconate, D-glucosaminic acid, D-glucose, glucose-1-phosphate, glucose-6-phosphate, glucuronamide, D-glucuronic

acid, glycyl-L-aspartic acid, α-hydroxy glutaric acid γ-lactone, *m*-inositol, D-mannitol, D-mannose, mucic acid, D-psicose, L-rhamnose, D-saccharic acid, D-sorbitol, tricarbalyllic acid; ($C_5$) adonitol, L-alanyl-glycine, L-arabinose, L-glutamate, L-glutamine, glycine betaine, α-keto butyric acid, α-keto glutarate, L-lyxose, mono-methyl succinic acid, L-proline, D-ribose, D-xylose, D-xylulose, L-xylulose; ($C_4$) acetoacetic acid, L-asparagine, D-aspartate, L-aspartate, bromo succinic acid, fumarate, α-hydroxy butyric acid, D-malate, L-malate, methyl pyruvate, oxaloacetic acid, succinate, *m*-tartaric acid, D-threonine, L-threonine; ($C_3$) D-alanine, L-alanine, glycerol, D,L-α-glycerol phosphate, L-lactic acid, malonate, propionic acid, pyruvic acid, 1,2-propanediol, D-serine, L-serine; ($C_2$) acetic acid, 2-amino ethanol, glycolate, glyoxylate; ($C_1$) formic acid.
Transport experiments making use of radiolabeled compounds were performed according to the rapid filtration method described elsewhere [123].

*Chromosomal deletion of the* yiaMNO *genes*

Upstream (F1) and downstream (F2) flanking regions were cloned as 1,620 bp PCR-fragments, introducing *Bam*HI-sites for fusion in pET908. The F1/F2-fragment was cloned into pORI240, creating pORIF1F2. Using this vector, an unmarked chromosomal deletion of the *yiaMNO* genes in *E. coli* strains MC4100 and JA134 was created as described elsewhere [114]. Cells were checked via PCR for deletion of the *yiaMNO* genes. The deletion mutants were labeled TP001 and TP018, respectively.

*Isolation of total RNA and RT-PCR*

Cells were grown in LB medium to an $OD_{660}$ of 1.0, harvested and frozen in liquid nitrogen. Lysozyme (20 μg/ml) was added to the pellets at room temperature (RT). Pellets were dissolved in 1 ml Trizol reagent (GIBCO BRL) and incubated at RT for 5 min. Chloroform (200 μl) was added, the suspension was mixed (> 15 sec), allowed to settle at RT for 3 min, and centrifuged (12,000 x g, 15 min, 4°C). The aqueous phase (500 μl) was transferred to a new tube, isopropanol (300 μl) was added and samples were briefly mixed. Samples were incubated at RT for 10 min and centrifuged as above. Pellets were washed twice with 70% (v/v) ice-cold ethanol (7,500 x g, 5 min, 4°C), air-dried and resuspended in a small volume of RNase-free demineralized water. Concentrations were determined using the Gene Quant system (Amersham Pharmacia Biotech), and samples were checked for DNA contamination via PCR, using primers detecting a 632-bp fragment of *secY*. RT-PCR was performed with RT-PCR-beads (Amersham Pharmacia Biotech) following manufacturer's instructions, using 1 μg total RNA per reaction. Primers detected a 200-bp internal fragment, except in the case of *yiaM* (474 bp complete fragment). Products were analyzed on ethidium bromide-stained 2% agarose gels.

**Table 3.** Primers used for PCR and RT-PCR

| Gene/fragment | Sequence | Forward/Reverse | Site[a] |
|---|---|---|---|
| Flanking region[b] 1 | CCCTGAATTCATTCAACAACTGGA | F | *Eco*RI |
| Flanking region[b] 1 | AAGTCGGATCCGCAGCAACCTG | R | *Bam*HI |
| Flanking region[b] 2 | AGGGATCCGATCTGCTGAAAGCC | F | *Bam*HI |
| Flanking region[b] 2 | CATATCTAGAAGCAGGGTGCGCA | R | *Xba*I |
| Deletion check | GATACCGGGCAGCTACGC | F | |
| Deletion check | ATACCGACAATTTGTTCCC | R | |
| *secY*, fragment[c] | GGCCTGGTGATTAACCCG | F | |
| *secY*, fragment[c] | CCGAATTCCTGGTACAAATGC TCCGGACTTCTT | R | |
| *yiaL*, fragment[d] | CTACCGAATTCAACGCCCTGGAGC | F | |
| *yiaL*, fragment[d] | TTGCGCGGATCCAATAGTGAT | R | |
| *yiaM* | TAGCCATGGAAAAAATACTCGAAGC | F | *Nco*I |
| *yiaM* | GGGGAATTCTTAAGCTCCTTGCGG | R | *Eco*RI |
| *yiaN*, fragment[d] | CCGATTGAATTCGGTGTCATG | F | |
| *yiaN*, fragment[d] | GCGGGATCCTTAATCCATTTCAAAGGG | R | |
| *yiaO* | GGGCCATGGAATTACGCTCTGTAACC | F | *Nco*I |
| *yiaO* | CCCGGATCCACCTTGCACCTCATCCAC | R | *Bam*HI |
| *yiaO*, fragment[d] | AAGTCCGGAATTCCAGCAGGC | F | |
| *yiaO*, fragment[d] | CCCTCTAGATTATTGCACCTCATCCAC | R | |
| *lyxK*, fragment[d] | AATACTGGATCCGGTTAGATTGTGG | F | |
| *lyxK*, fragment[d] | CCGAGAATTCGTTCCCCGCTAACA | R | |
| *yiaK*, promoter | TCCCCATTTTGTCGCGTCCTG | F | |
| *yiaK*, promoter | ACGCCGCGTGAAATTAAGAC | R | |

[a] restriction endonuclease site introduced by primer

[b] 1,620 bp up- (1) and downstream (2) regions flanking *yiaMNO*

[c] 632-bp internal fragment

[d] 200-bp internal fragment

*Construction of the* yiaK-lacZ *transcriptional fusion*

The 268-bp 5'-upstream region of the *yiaK* gene was amplified by PCR and transcriptionally fused to *lacZ* by insertion into plasmid pRS550 [191]. This plasmid carries a promoter-less *lac* operon and genes that confer resistance both to Km and Am. Recombinant plasmids were selected, after transformation of strain XL1-Blue, as blue colonies on LB plates containing X-gal, Km and Am. Correct orientation of the inserted fragment was confirmed by sequencing of the plasmid DNA using an M13 primer. Merodiploids were obtained by transferring the fusions as single copies into the *trp* operon of *E. coli* strain TE2680 as described [56]. Transformants were selected for Km resistance and screened for sensitivity to Am and chloramphenicol. P1 *vir* lysates were made to transduce the fusions into strain ECL1.

*Tetrazolium violet metabolic assays*

Prefabricated ES Microplates (Biolog) were used according to the manufacturer's instructions to test carbon source usage by strains MC4100 and TP001. Metabolic properties of strains ECL1, JA134, and TP018 were investigated using the redox indicator tetrazolium violet (Sigma) [24]. Assays were performed in sterile 96-wells microplates (Greiner) in a total volume of 150 μl per well. Tetrazolium violet and carbon sources were added to final concentrations of 0.0025 % and 0.2 % (w/v), respectively. Cells were scraped from a LB-agar plate and resuspended in M63 minimal medium, added to the pre-filled plates, and incubated overnight at 37ºC. Tetrazolium violet reduction was analyzed by measuring absorbance at 590 nm using a Spectramax 340 titertek-reader (Molecular Devices).

*Proteinase protection assay*

Purified YiaO-His (see below) was dialyzed extensively against 50 mM potassium phosphate, pH 7.4; 100 mM NaCl; 2 mM EDTA at 4°C. 1 mg/ml YiaO-His (27 μM) was incubated with 4 mM substrate at RT for 10 min, transferred to a Proteinase K (Merck) solution (10 μg/ml final concentration) and incubated for 15 min on ice. The reaction was stopped by boiling the samples in SDS-PAGE sample buffer. Samples were analyzed on 15% SDS-PAGE gels followed by CBB staining.

*Enzymatic detection of L-xylulose*

NADP-xylitol dehydrogenase catalyzes the stereospecific reaction [76]:

$$\text{L-xylulose} + \text{NADPH} + H^+ \rightarrow \text{xylitol} + NADP^+$$

Utilization of L-xylulose was monitored by recording the decrease in absorbance at 340 nm due to consumption of NADPH, which is proportional to the L-xylulose concentration [14]. NADP-xylitol dehydrogenase was isolated from 5 grams of acetone-dried guinea pig liver (Sigma) by a modification of

the method of Ashwell [14], and finally dissolved in 5 ml demineralized water. Aliquots of the enzyme solution were frozen in liquid nitrogen and stored at -80°C until further use. In the assay, to a final volume of 1800 µl were successively added: 75 mM Tris-HCl, pH 7.0; 5 mM $MgCl_2$; 1 mM cysteine-HCl; 100 µM β-NADPH; 50-100 µl enzyme suspension. The signal was allowed to stabilize, before 200 µl sample with unknown L-xylulose content was added. Measurements were performed at 37ºC under continuous stirring, using a spectrophotometer (Cary). Values were compared with a calibration curve, which was linear in the range of 10-60 µM [14].

*Overexpression and purification of YiaO*

The *yiaO* gene was cloned via PCR and ligated into pSA5, yielding vector pET917 encoding YiaO with a carboxyl-terminal 6xhistidine tag. *E. coli* SF100 was transformed with pET917, grown to an $OD_{660}$ of 0.6, and overexpression was induced by addition of 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). Cells were harvested (4,000 x g, 10 min, 4ºC), and the periplasmic fraction was isolated using the cold osmotic shock procedure [148]. The shock fluid was supplemented with 50 mM potassium phosphate, pH 7.4; 100 mM NaCl; 15 mM imidazole, and incubated overnight with $Ni^{2+}$-NTA agarose (Qiagen) at 4ºC. The suspension was poured into a column and unbound material was eluted. The column was washed twice with buffer (50 mM potassium phosphate, pH 7.4; 100 mM NaCl; 15 mM imidazole), and YiaO-His was eluted in this buffer containing 200 mM imidazole. Fractions were analyzed on 15% SDS-PAGE, followed by Coomassie Brilliant Blue (CBB) and silver staining.

*Substrate binding assay*

L-xylulose binding by YiaO-His was studied by modifying a substrate retention assay described elsewhere [206]. 100 µM purified YiaO-His was incubated with 50 µM L-xylulose for 10 min at 37°C. BSA was used at the same concentration. D-glucose and D-xylulose were used at 1 mM final concentration. Protein was removed by centrifugation (10,000 x g, 45 min, 4°C) using Microcon YM-10 (10 kDa cut-off) filters. The flow through was analyzed using the L-xylulose detection assay.

*Utilization of L-xylulose*

Strains ECL1, JA134 and TP018 were grown in LB to an $OD_{660}$ of 1.0, and harvested as above. For whole cell uptake experiments, cells were resuspended to an $OD_{660}$ of 10 in M63 medium, and diluted 1:1 with M63 containing L-xylulose (0.5 mM final concentration). Cell suspensions were incubated at 37°C under continuous shaking, and 250 µl-samples were taken at various intervals. Cells were removed by centrifugation (13,000 x g, 5 min, RT) and 200 µl of the supernatant was analyzed for L-xylulose content. Spheroplasts were prepared as described by Cao *et al.* [32]. 300 µl spheroplasts (10 mg/ml final concentration) in 50 mM potassium phosphate, pH 7.4; 5 mM $MgSO_4$, were mixed with 100 µl YiaO-His (1 mg/ml), and pre-incubated with 1.2 mM L-xylulose for 5 min at RT. The suspension was shaken at 37°C

for 30 min, followed by centrifugation (13,000 x g, 5 min, RT), and the L-xylulose content of the supernatants was determined.

*Miscellaneous methods*

β-galactosidase activity was assayed by hydrolysis of *o*-nitrophenyl-β-D-galactopyranoside (ONPG) and expressed as Miller Units [142]. Protein content was determined using the DC protein assay (Biorad).

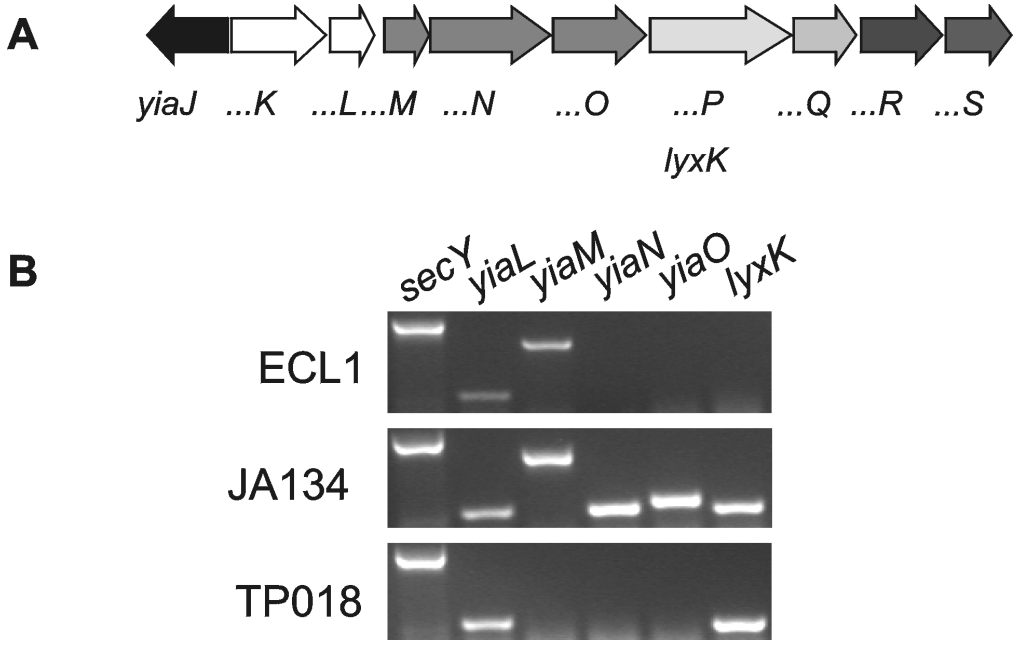



**Figure 1** (A) Structural organisation of the *yiaJ-S* gene cluster of E. coli K-12. *yiaJ*: putative regulator; *yiaK*: putative dehydrogenase; *yiaL*: unknown function; *yiaM*: small membrane domain of transporter (4 TMDs); *yiaN*: large membrane domain of transporter (12 TMDs); *yiaO*: periplasmic substrate-binding protein; *yiaP* (*lyxK*): kinase, phosphorylates both L-xylulose and 3-keto-L-gulonate; *yiaQ*: 3-keto-L-gulonate 6-phosphate decarboxylase (also: putative hexulose-6-phosphate synthase); *yiaR*: putative hexulose-6-phosphate isomerase; *yiaS*: ribulose-5-phosphate 4-epimerase [22, 87, 88, 244].
(B) Deletion of the *yiaMNO* structural genes does not affect expression of up- (*yiaL*) and downstream (*lyxK*) genes. Expression of the *yiaL-lyxK* region in the three strains was detected via RT-PCR. Primers were designed to detect mRNA fragments of *secY* (control for constitutive expression, lane 1), *yiaL* (lane 2), *yiaMNO* (the putative transporter, lanes 3-5), and *lyxK* (lane 6).

## Results

*Deletion of the* yiaMNO *genes*

Strain JA134 constitutively expresses the *yiaK-S* cluster (Fig. 1A) [87, 183]. To investigate the role of the YiaMNO transporter, an unmarked chromosomal deletion of the three structural genes was constructed in this strain, yielding strain TP018. The effect of this deletion on expression of the genes encoded at this locus was studied by RT-PCR, using primers directed against *yiaL*, *yiaM*, *yiaN*, *yiaO*, *lyxK*, and *secY* (control). Only low level expression of the *yiaL* and *yiaM* genes was detected in the JA134 parental strain ECL1 (Fig. 1B). None of the other genes were expressed in this strain under the conditions of the assay. In strain JA134 all five genes were strongly expressed (Fig. 1B) which is in agreement with previous observations [87]. Deletion of *yiaMNO* did not affect transcription of the genes located immediately up- and downstream (Fig. 1B, TP018).

*Carbon source metabolism by strains ECL1, JA134 and TP018*

Strains ECL1, JA134 and TP018 were analyzed for their ability to metabolize potential substrates of the YiaMNO transporter, using the irreversible reduction of tetrazolium violet to its purple formazan as an indicator for carbon source catabolism [24]. L-lyxose was metabolized by strains expressing the cluster, but deletion of the *yiaMNO* genes had no significant effect (Fig. 2, error bars). L-lyxose is a substrate for the L-rhamnose-$H^+$ symporter RhaT [16, 145] and *rhaT* deletion mutants of JA134 do not grow on L-lyxose (J.Badia and J. Aguilar, unpublished results). Thus, L-lyxose enters the cell via RhaT, and not via the YiaK-S system. A recent report showed that the kinase LyxK phosphorylates 3-keto-L-gulonate, a breakdown product of L-ascorbate metabolism or the result of the reduction of 2,3-diketo-L-gulonate [244]. L-ascorbate is a substrate for cells that constitutively express the cluster, but metabolism of this compound is not affected in strain TP018 (Fig. 2). 2,3-diketo-L-gulonate, which is not commercially available, was not tested in these experiments, but its presence in L-ascorbate solutions at pH 7.0 has been reported [192].

However, no induction of the *yiaK-S* cluster by L-ascorbate-derived 2,3-diketo-L-gulonate, making use of the ϕ(*yiaK-lacZ*) in the genetic background of strain ECL1, was detected (data not shown). Therefore, 2,3-diketo-L-gulonate probably is not a substrate of the *yiaK-S* operon.

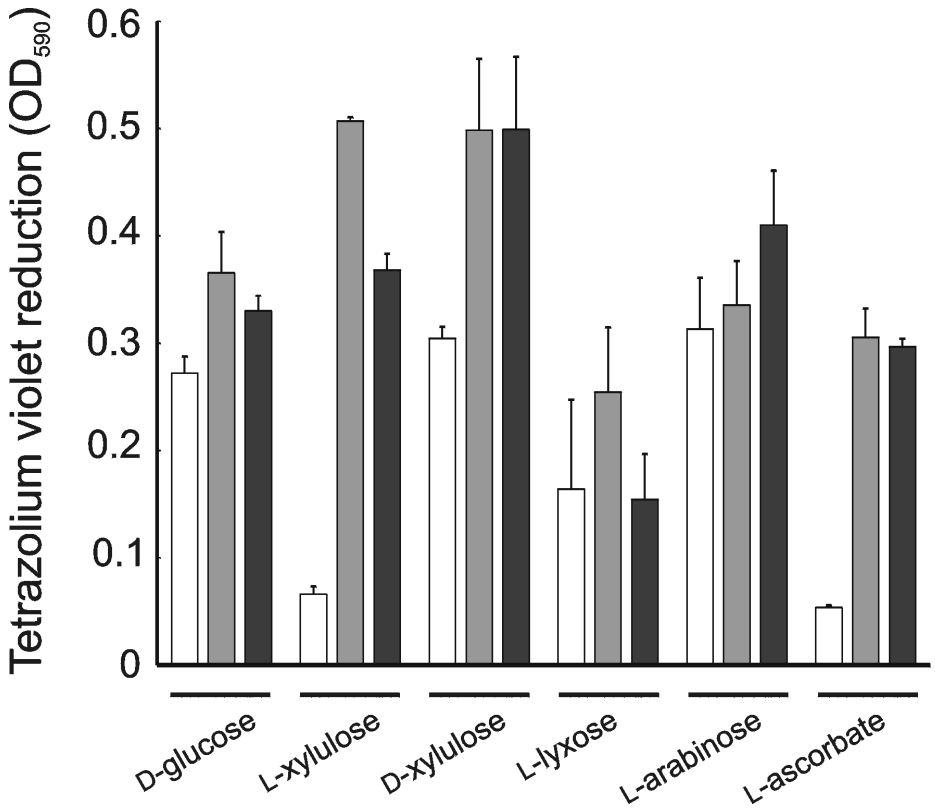



**Figure 2.** Only cells expressing the *yiaK-S* gene cluster metabolize L-xylulose, and deletion of the YiaMNO transporter negatively affects this property. Metabolic activity of the three strains on various putative carbon sources was measured via tetrazolium violet reduction, which was visualized through formation of its purple formazan (expressed in $OD_{590}$ values). ECL1: white bars; JA134: grey bars; TP018: black bars.

Expression of the *yiaK-S* cluster allowed JA134 to utilize L-xylulose, whereas ECL1 did not (Fig. 2). Effects of the *yiaMNO* deletion were only observed on this pentose, as TP018 metabolized L-xylulose less efficiently than JA134 (Fig. 2). However, deletion of the *yiaMNO* genes did not completely abolish L-xylulose metabolism, therefore an additional L-xylulose-transporting system exists. Further work focused on L-xylulose as a potential substrate for the YiaMNO transporter.

*L-xylulose is taken up by strains JA134 and TP018*

Solute transport studies are most conveniently performed with radioactively labeled substrates. However, L-xylulose is not available in radiolabeled form, and due to its physiochemical properties it is also not possible to tritiate this compound. Therefore, we have made use of assays that allow measurement of L-xylulose consumption by whole cells in a non-radioactive manner.

NADP-xylitol dehydrogenase is a stereo-specific enzyme that catalyzes the reduction of L-xylulose to xylitol coupled to the oxidation of NADPH to $NADP^+$ (see 'Experimental procedures') [76]. The consumption of NADPH can be monitored at 340 nm, and is directly proportional to the L-xylulose content of the mixture [14]. Strains ECL1, JA134 and TP018 were incubated in M63 medium in the presence of L-xylulose as the sole carbon and energy source. Using the enzyme assay, the uptake of L-xylulose from the extracellular medium by the cells was monitored in time. Wild type strain ECL1 did not consume the pentose (Fig. 3). In contrast, strain JA134 rapidly metabolized L-xylulose (Fig. 3). The *yiaMNO* deletion strain TP018 was still capable of utilizing L-xylulose, but reproducibly at a lower initial rate than strain JA134 (Fig. 3). These findings confirm that L-xylulose is taken up and metabolized by strains that constitutively express *yiaK-S*, and the YiaMNO transporter is involved in, but not essential for L-xylulose utilization in this genetic background.

*L-xylulose partially protects YiaO from protease digestion*

Solute-binding proteins undergo a large conformational change upon substrate binding, which often is accompanied by a changed susceptibility towards proteolytic degradation. The ability of L-xylulose to protect YiaO against degradation by Proteinase K was investigated. YiaO was overexpressed with a C-terminal His-tag and purified to homogeneity (Fig. 4, 'no additions'), pre-incubated with an excess of substrate, and subjected to Proteinase K-mediated degradation. Both in the absence of substrate and in the presence of D-glucose, the protein was nearly completely degraded (Fig. 4). L-xylulose protected YiaO-His, as did D-xylulose to some extent (Fig. 4), indicating that both pentoses can interact with YiaO-His.

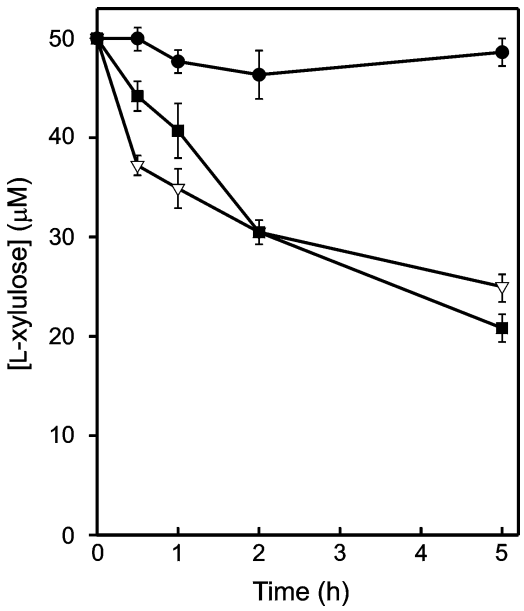



**Figure 3.** Uptake of L-xylulose by whole cells. The extracellular L-xylulose concentration was monitored in time for strains ECL1 (black circles), JA134 (white triangles), and TP018 (black squares). ECL1 is unable to metabolize L-xylulose, while strain JA134 rapidly consumes the pentose. The initial uptake (t = 0-1h) of TP018 lags behind that of JA134, but eventually both activities are nearly indistinguishable (t > 1h).

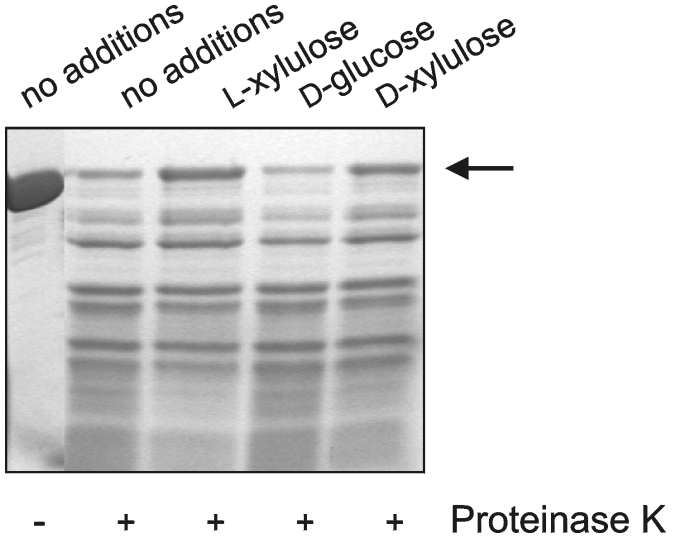



**Figure 4.** Protection against Proteinase K digestion of YiaO by addition of substrate. Both L- and D-xylulose protect purified YiaO ('no additions'; no Proteinase K) to some extent from proteolysis, while D-glucose does not.

*YiaO-His binds L-xylulose and stimulates metabolism by spheroplasts expressing* yiaMN

To verify that YiaO-His binds L-xylulose, we modified the non-radioactive substrate-binding assay of Tetsch and Kunte [206]. Purified YiaO-His was incubated with L-xylulose and then filtered from the solution, and the L-xylulose-content of the flow through was determined using the enzyme-based reaction. Approximately 10% of the L-xylulose had been removed from the supernatant as a result of binding by YiaO-His (Fig. 5). In contrast, the pentose remained in solution in the absence of protein, or in the presence of bovine serum albumin (BSA, Fig. 5) which served as a control for non-specific binding to protein. Binding of L-xylulose was not competed by a 20-fold excess of either D-glucose or D-xylulose (Fig. 5). Although this method is unsuitable for estimating the dissociation constant for binding, the data indicate that YiaO-His is able to bind L-xylulose. Since D-xylulose did protect against Proteinase K mediated proteolysis but does not compete for L-xylulose binding in the retention assay, YiaO binds this pentose most likely with lower affinity.

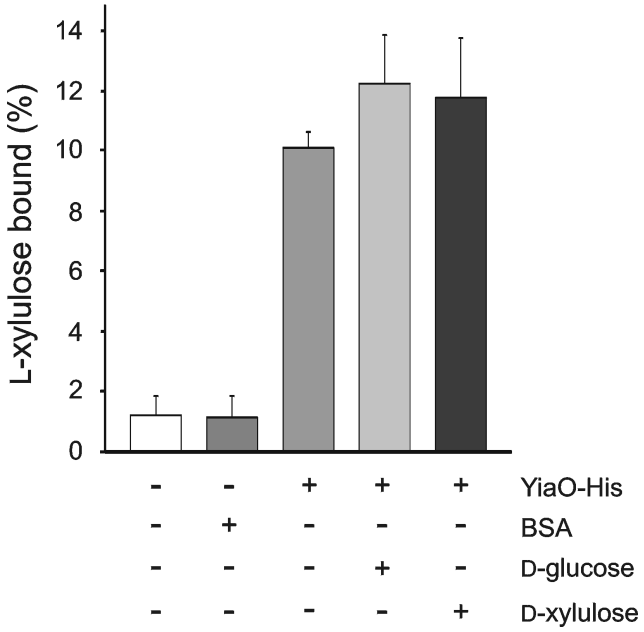



**Figure 5.** Purified YiaO-His specifically binds L-xylulose. Percentage of L-xylulose bound by YiaO-His in the retention assay, in the presence or absence of an excess of the potential competitors D-glucose and D-xylulose. BSA: control for non-specific binding to protein.

Spheroplasts prepared from strains ECL1, JA134, and TP018 were incubated with L-xylulose in the absence and presence of purified YiaO-His. The concentration of the L-xylulose remaining in the supernatant was determined after incubation. ECL1 spheroplasts did not metabolize L-xylulose (Fig. 5). Spheroplasts of strains JA134 and TP018 both consumed L-xylulose, but stimulation of L-xylulose utilization was observed only upon addition of purified YiaO-His to JA134, but not TP018, spheroplasts (Fig. 5). Taken together, these experiments show that YiaO-His is able to bind L-xylulose, and together with the YiaMN proteins can function as a binding protein-dependent uptake system for L-xylulose.

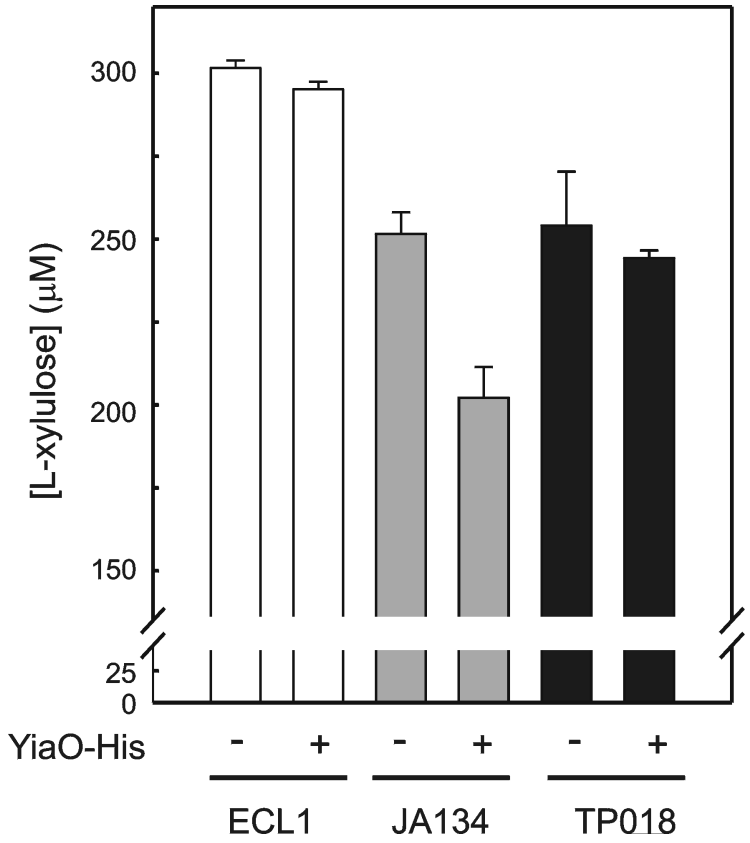



**Figure 6.** YiaO-His stimulates L-xylulose uptake by *yiaMN*-expressing spheroplasts. L-xylulose uptake by spheroplasts prepared from the three strains, in the absence (-) and presence (+) of YiaO-His, was determined by measuring the concentration of the pentose in the supernatant after incubation. ECL1 (white bars) does not take up the pentose. Both JA134 (grey bars) and TP018 (black bars) spheroplasts consume L-xylulose but only the JA134 spheroplasts, that express *yiaM* and *yiaN*, are stimulated by the addition of YiaO-His.

## Discussion

The *yiaMNO* genes of *E. coli* K-12 encode a putative binding protein-dependent secondary transporter [53, 59, 98, 171]. In order to investigate the role of the YiaMNO transporter, we have performed an elaborate screen to identify the substrate using *E. coli* K-12 strain MC4100 and its Δ*yiaMNO* derivative (T.H. Plantinga and D. Tomkiewicz, unpublished results). Over 100 compounds were tested (see 'Experimental procedures'), but none yielded a positive result. However, using strain JA134, we were able to identify one substrate for the transporter. Whole cells utilized L-xylulose only when the *yiaK-S* cluster was constitutively expressed, and the initial L-xylulose metabolic activity was lowered when the YiaMNO transporter was deleted (Figures 2 and 3). Interestingly, the data show that there is a second system present in these *E. coli* strains that can facilitate uptake of L-xylulose, but this is only detected in cells that express the metabolizing enzymes (see below).

L-xylulose transport has not been studied before, mainly because this compound is not available in radiolabeled form. Evidently, this also hampered the analysis presented in this report. Therefore, indirect binding and transport assays had to be developed. The data show that YiaO, the periplasmic substrate-binding protein of the transporter, specifically binds L-xylulose (Figures 4 and 5). Addition of YiaO-His to spheroplasts stimulated metabolism of the pentose only when the membrane domains YiaMN were expressed (Fig. 6). Taken together, these findings point to a role of the YiaMNO transporter in L-xylulose uptake. However, the YiaMNO system is neither induced by L-xylulose (J. Badia and J. Aguilar, unpublished results), nor do wild type cells grow on L-xylulose. Therefore, the physiological function of L-xylulose uptake by *E. coli* K-12 is unclear. Since the pentose is not abundant in nature it may well be that the transporter is involved in uptake of other pentoses.

In *E. coli* K-12, the *yiaK-S* genes are tightly regulated [17, 87, 183], but have not been inactivated or lost during evolution. Instead, expression is repressed by YiaJ, a member of the IclR family of transcriptional regulators [17, 87]. Recently, the structure of a protein homologous to YiaJ from *Thermotoga maritima* has been solved [247]. Although a binding site for an inducer could be assigned, the identity of this compound

is unknown, and the physiological conditions where this system is induced remain to be identified.

The *yiaMNO* deletion only partially abolishes L-xylulose metabolism. This suggests that an additional transport system must be present that is capable of L-xylulose uptake. The second system is most likely also present in the wild type strain ECL1, but its activity escapes detection due to the absence of the L-xylulose-metabolizing enzymes encoded downstream of the transporter in the *yiaK-S* cluster. Therefore, it is less likely that the second uptake system is strictly dedicated to the uptake of L-xylulose. There are several L-sugar uptake systems described for *E. coli* K-12 that may be involved in this uptake activity. The L-rhamnose-$H^+$ symporter RhaT transports the structurally related sugar L-lyxose [145, 205], and the L-fucose-$H^+$ symporter FucP can also transport L-galactose and D-arabinose [68, 145]. Neither of these systems has been tested for L-xylulose transport. However, the specificity of these transporters is determined by the nature of the side-chains at both the C-2 and the C-4 positions of the pyranose rings [145], while L-xylulose forms a furanose ring. L-xylulose does not compete with L-rhamnose for transport by RhaT (J. Badia and J. Aguilar, unpublished results).

The mechanistic properties of the YiaMNO transporter need further investigation, but since there is no radiolabeled L-xylulose available these studies are extremely difficult to perform at this time. In addition, the physiological function of L-xylulose uptake and metabolism by *E. coli* K-12 is unclear. Binding protein-dependent secondary transporters have previously been implicated in the uptake of organic anions and compatible solutes. Our report extends the substrate range to a pentose sugar.

## Acknowledgements

We would like to thank Sonja Albers for plasmid pSA5, and Danka Tomkiewicz for valuable practical assistance. This work was supported by the Netherlands Organization for Scientific Research (N.W.O., grant 805-19-046 P) and by the Dirección General de Investigación, Ministerio de Ciencia y Tecnología, Madrid, Spain (grant BMC2001-3003).

# *Chapter 5*

## Different transporter classes genetically linked to a conserved carbohydrate metabolic cluster: exemplary for transporter evolution?

*"It's evolution, baby!"* [219]

Titia H. Plantinga, Chris van der Does and Arnold J. M. Driessen

(This chapter has been accepted for publication in *Trends in Microbiology*)

**Summary**

The yiaQRS genes of Escherichia coli K-12 are involved in carbohydrate metabolism. Clustering of homologous genes was found throughout a number of unrelated bacteria. Strikingly, all four bacterial transport protein classes were found, conserving transport function but not mechanism. It appears that during evolution the ability the transport, phosphorylate, and metabolize the substrate have been conserved in these operons, but the transporter classes were swapped. This probably demonstrates the subtlety of transport protein evolution.

## Introduction

The gene cluster designated *yiaJKLMNOPQRS* (GenBank accession number g1789999-1790008) in the *Escherichia coli* K-12 genome (Fig. 1a) has been suggested to be involved in the uptake and metabolism of carbohydrates [17, 87, 88, 173, 183]. Recently, we have shown that the binding protein-dependent secondary transporter encoded by the *yiaMNO* genes transports the rare pentose L-xylulose [Chapter 4]. During characterization of this transporter we became interested in the function of the remaining genes in the cluster. The increase in available genome databases allowed us to search for the presence of similar gene clusters. Homologs of the individual *yiaQ*, *yiaR* and *yiaS* genes were found throughout bacterial and archaeal genomes (data not shown). Remarkably, the clustering of these genes in putative operons occurs in a limited number of bacterial genomes. The genomic organisation suggests that these genes specify similar functions.

## Clustering of putative carbohydrate metabolizing enzymes

An extensive search of accessible genome databases discovered 30 gene clusters containing putative enzymes that are highly similar to the *yiaQRS* genes, and in most cases the gene order is very well conserved (Fig. 1: grey). Interestingly, representatives are found in several unrelated bacterial species, both Gram-negative and Gram-positive, most of which are known human pathogens (Table 1). In fact, clustering of these three enzymes is found in 26 of the (completely) sequenced bacterial genomes available at this time, even in the relatively small genomes of three *Mycoplasma* species (Table 1). Although homologues of these genes are found in archaea, the cluster seems not conserved in this kingdom.

Up to now, only the *E. coli* K-12 enzymes have been biochemically characterized to some detail. YiaQ was shown to function as a 3-keto-L-gulonate 6-phosphate decarboxylase in a comparison with its homolog YjfV (see below) [244]. Based on amino acid sequence homology YiaQ has previously been suggested to function as a hexulose 6-phosphate synthase [173], but this activity could not be demonstrated experimentally [243].

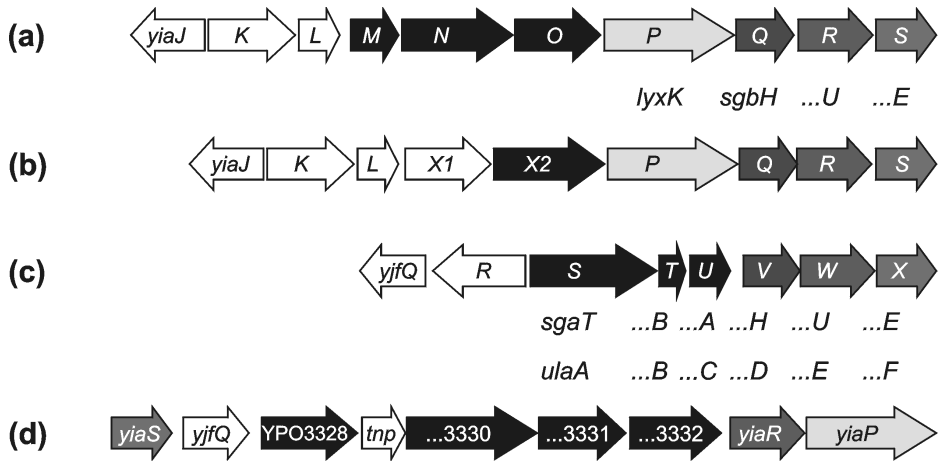



**Figure 1.** Structural organization of highly similar gene clusters containing *yiaQRS* homologs (grey; see text for functional details) that cluster with members from either of the four classes of transport proteins (black) as found in four model organisms. Conventional nomenclature is used, with proposed alternatives given below each gene. Identical shades of grey: genes encode homologous enzymes. Light grey: (putative) kinases. **(a)** The *yiaJ-S* gene cluster of *E. coli* K-12 contains a binding protein-dependent secondary transporter (*yiaMNO*). **(b)** The *yiaJ-S* gene cluster of *K. oxytoca* contains a secondary transporter (*yiaX2*). **(c)** The *yjfQ-X* gene cluster of *E. coli* K-12 contains a PTS system (*yjfS*: permease IIC component, *yjfT* and *yjfU*: cytosolic phosphotransferase components IIB and IIA, respectively). **(d)** The YPO3328-3332 genes of *Y. pestis* encode a binding protein-dependent primary (ABC) transporter (YPO3328: periplasmic substrate-binding protein, YPO3330: ATPase domain, YPO3331 and YPO3332: permease domains). The gene-order is not conserved, but *yiaR*, *yiaS*, and *yiaP* homologs are present. *yiaJ* and *yjfQ*: putative regulators; *tnp*: transposase. [17, 31, 87, 88, 183, 244, 248]

The YiaR amino acid sequence shows a high similarity to 3-epimerases, and has therefore been tentatively labeled as a L-xylulose 5-phosphate 3-epimerase [88]. YiaS was demonstrated to be a ribulose 5-phosphate 4-epimerase [88, 244]. The *yjfVWX* genes of *E. coli* K-12 (Fig. 1c) are highly similar in sequence to the *yiaQRS* genes [173, 244]. The YjfV protein is 46% identical to YiaQ, and has been shown to function as a 3-keto-L-gulonate 6-phosphate decarboxylase [244]. YjfW is a L-xylulose 5-phosphate 3-epimerase, and is 56% identical to YiaR [244]. Finally, YjfX is a ribulose 5-phosphate 4-epimerase, and 61% identical to YiaS [244]. The *yjfQ-X* cluster of *E. coli* K-12 (Fig. 1c) has recently been shown to facilitate the uptake and subsequent metabolism of L-ascorbate [244, 248].

Detailed functional information concerning the gene clusters discovered in our search is lacking at this time, and the ground for their strong conservation in a range of potential human pathogens is unknown. This intriguing question awaits further investigation.

**Conservation of transport function**

The enzyme clusters have been described before for their potential roles in carbohydrate metabolism [173], but an intriguing feature of these systems has been overlooked up to now. It appears that the clustering of the genes encoding these enzymes in all cases coincides with the presence of a transport system (Fig. 1: black). Strikingly, the type of transporter varies (Table 1; Fig. 1). Each of the four classes of prokaryotic solute uptake systems known to date (see Fig. 2) is found, and representative gene clusters are described here.

The most widespread cluster (20 out of 30), and the predominant one found in Gram-positive bacteria, contains a PTS (phospho*enol*pyruvate:carbohydrate phosphotransferase system) [165, 173] (Fig. 1c: '*yjfSTU*'; and Fig. 2c). PTS systems employ multiple cytosolic proteins, such as the IIA and IIB components, to transfer the phosphoryl group from phospho*enol*pyruvate (PEP) to the substrate that is taken up via the permease IIC component (Fig. 1c; and Fig. 2c) [165]. These systems are found only in bacteria and mainly take up carbohydrates, although the YjfSTU system of *E. coli* was recently shown to transport L-ascorbate [248]. Secondly, the *yiaJ-S* cluster of *E. coli* K-12 contains a binding protein-dependent secondary transporter [53, 54], or tri-partite ATP-independent periplasmic (TRAP) [98, 171] transporter (Fig. 1a: '*yiaMNO*'; and Fig. 2a). In this type of transporter, the unidirectional uptake of solutes depends on the solute-binding protein, which delivers the substrate to the two membrane-localized permease domains from the extracellular side (Fig. 2a). Transport is driven by the proton- (pmf) and/or sodium- (smf) motive force. The membrane proteins are dissimilar: the small permease component (e.g. YiaM) contains 4 putative transmembrane domains (TMDs) and shows no homology with known transport proteins. The large domain (e.g. YiaN) resembles the classical secondary transporters (see below), i.e. it contains 12 putative TMDs and a large central cytosolic loop that separates TMD 6 from TMD 7 [98, 53, 171].

**Table 1.** Bacterial genomes containing the putative carbohydrate metabolism cluster

| Organism and strain: | Locus: | Transporter: | | | |
|---|---|---|---|---|---|
| | | BPD sec[a] | Secondary[b] | BPD ABC[c] | PTS[d] |
| **Gram-negative:** | | | | | |
| *Escherichia coli* K-12 (MG1655) | AE000435 | + | - | - | + |
| *E. coli* O157:H7 (EDL933 and Sakai) | AE005174/BA000007 | - | - | - | + |
| *E. coli* CFT073 | AE016768 | + | - | - | + |
| *Salmonella typhimurium* LT2 | AE008870 | + | - | - | + |
| *S. enterica* subsp. *enterica* serovar *Typhi* (CT18) | AL513382 | +[e] | - | - | - |
| *Klebsiella oxytoca* | AF282849 | - | + | - | - |
| *Shigella flexneri* (2a str. 301) | AE015441 | - | - | - | + |
| *Yersinia pestis* (CO92) | AL590842 | - | - | +[f] | - |
| *Haemophilus influenzae* Rd | L42023 | + | - | - | - |
| *H. somnus* (129 PT) | NZ_AABO02000002 | + | - | - | - |
| *Pasteurella multocida* | AE004439 | +[f] | - | - | +[g] |
| *Vibrio cholerae* El Tor | AE003853 | - | - | - | +[h] |
| *V. vulnificus* (CMCP6) | AE016811 | - | - | - | + |
| **Gram-positive:** | | | | | |
| *Mycoplasma penetrans* | AP004173 | - | - | - | + |
| *M. pneumoniae* (M129) | U00089 | - | - | - | +[i] |
| *M. pulmonis* | AL445565 | - | - | - | +[i] |
| *Streptococcus agalactiae* (2603V/R) | AE014274 | - | - | - | +[i] |
| *S. mutans* | AF397165 | - | - | - | +[i] |
| *S. pneumoniae* (R6 and TIGR4) | AE007317/AE005672 | - | - | - | +[i] |
| *S. pyogenes* (M1 GAS and MGAS8232) | AE004092/AE009949 | - | - | - | +[i] |
| *Enterococcus faecium* | NZ_AAAK01000214 | - | - | - | +[i] |
| *Leuconostoc mesenteroides* subsp. *mesenteroides* | NZ_AAABH02000028 | - | - | - | + |
| *Oceanobacillus iheyensis* (HTE831) | AP004602 | + | - | - | - |

[a] binding protein-dependent secondary transporter
[b] secondary transporter
[c] binding protein-dependent ABC transporter
[d] PTS system
[e] *yiaO* interrupted by IS-element
[f] mixed gene order
[g] no *yjfVWX* homologues
[h] no IIB component, mixed gene order
[i] no *yjfQR*

Remarkably, a nearly identical cluster was found in *Klebsiella oxytoca*, but here the three genes encoding the transporter appear to have been exchanged for two different genes: one encoding a secondary transporter (Fig. 1b: '*yiaX2*'; and Fig. 2b), and one gene of unknown function with homology to a chemotaxis protein (Fig. 1B: '*yiaX1*').

Secondary transporters facilitate solute uptake in a pmf- and/or smf-dependent manner [163]. Transport is reversible, and is mediated by a single integral membrane protein that generally consists of 12 TMDs, with a large cytoplasmic loop between TMD 6 and TMD 7 [163] (Fig. 2b). Finally, also a binding protein-dependent ATP-binding cassette (ABC) transporter [78] (Fig. 2d) has been found. In *Yersinia pestis*, a *yiaR* and a *yiaS* homologue are located in a gene cluster with such a transport system (Fig. 1d: 'YPO3328-3332'). In these transporters, the solute-binding protein donates the solute to two identical or homologous membrane-localized transport domains. Uptake is driven by hydrolysis of ATP, which is mediated by two identical or homologous ATPases that are associated with the carrier at the cytosolic side of the membrane [78] (Fig. 2d). The *Yersinia* cluster also contains a kinase homologous to *yiaP* (see below), and a DeoR-like regulator (Fig. 1d: '*yjfQ*') that is homologous to the one located upstream of the YjfSTU PTS system of *E. coli* K-12 [31, 248]. Although the gene order at this locus is not conserved, this localization suggests these genes may specify similar functions.

**Conservation of phosphorylation function**

Notably, in all cases, the capacity to phosphorylate the substrate of the transporter has been conserved. The most widespread PTS system phosphorylates its carbohydrate substrate during transport [165] (Fig. 2c). Directly behind the three other types of transporters, a kinase that is able to phosporylate the substrate is found (Fig. 1: light grey). For example, the YiaMNO transporter of E. coli K-12 transports L-xylulose [Chapter 4], while the *yiaP* gene of *E. coli* K-12 encodes a L-xylulose kinase [183]. It is therefore highly likely that these systems are involved in the uptake and metabolism of a carbohydrate.

**Conclusion**

Taken together, 26 of those bacterial genomes that are available contain a highly similar gene cluster that probably specifies functions in carbohydrate metabolism. Both function and gene-order are remarkably conserved indicating that strong selection pressures are keeping these gene clusters together [201].

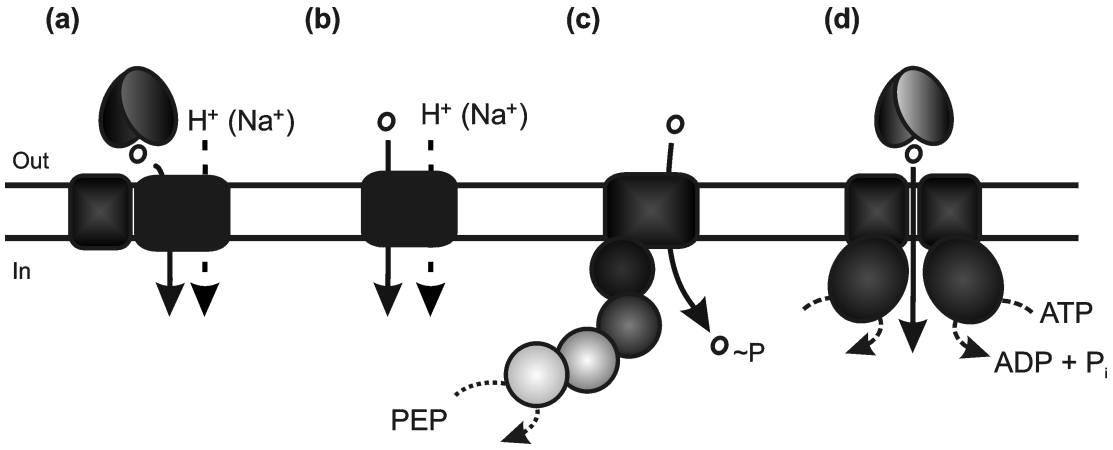



**Figure 2.** Domain structure of the four major classes of bacterial uptake systems (see text for details). **(a)** Binding protein-dependent secondary transporters [53, 54, 98, 171]. **(b)** Secondary transporters [163]. **(c)** Phospho*enol*pyruvate:carbohydrate phosphotransferase systems (PTS) [165]. **(d)** Binding protein-dependent primary transporters [78].

Strikingly, these clusters always contain a member from one of the four major classes of prokaryotic solute uptake systems. Transport function has thereby been conserved, but the mechanism by which the substrate is taken up varies. These transport systems are most likely responsible for the uptake of a carbohydrate, as has recently been demonstrated for the YiaMNO transporter of *E. coli* K-12 that transports L-xylulose [Chapter 4]. Although different transport mechanisms are used, the ability to phosphorylate the substrate is maintained. Following transport and phosphorylation, the substrate is further metabolized by the enzymes encoded downstream in the cluster.

Finally, the finding of the *K. oxytoca* gene cluster (Figure 1b), where transporter genes located at a central position appear to have been exchanged compared to the homologous *E. coli* K-12 cluster (Figure 1a), demonstrates the subtlety by which evolution can take place at the genome level. The functional description of these intriguing gene clusters awaits further investigation.

**Acknowledgement**

We thank Wil Konings for helpful discussion.

# *Chapter 6: General Discussion*

*"It's only a model!"* [156]

**Binding protein-dependent secondary transport…**

Although the binding protein-dependent secondary transporters were described in 1996, to this day only a few these transporters have been functionally characterized. At the start of the research presented in this thesis, only two members of this transport protein family had been functionally described. It had been shown that *Rhodobacter sphaeroides* expresses a binding protein-dependent uptake system for glutamate that is sensitive to agents that collapse the proton-motive force (pmf) [91]. The system is expressed when millimolar amounts of $Na^+$ are added to the growth medium, and allows the cells to utilize glutamate as a carbon source. Although the role of $Na^+$ is unclear, the ion is required for transport but not for binding of the substrate by the substrate-binding protein [91]. No sequence information is available that would identify this system in the currently accessible genome databases. A second family member that was identified is the DctPQM transporter of the related bacterium *Rhodobacter capsulatus* [59]. DctPQM transports the following $C_4$-dicarboxylates with decreasing affinity: L-malate > fumarate > succinate > D-malate [232], and allows the bacterium to use these compounds as carbon sources during aerobic growth [59]. Transport was inhibited by uncouplers and dependent on the presence of an extracellular binding protein [59, 232].

An extensive search of the available (partially) sequenced genomes, using the DctPQM amino acid sequences as query, revealed the presence of homologous systems in several unrelated bacteria [53, 59], including *Escherichia coli* K-12. By now it has been realized that the family is even more widespread, with representatives also in the Archaeal Kingdom [98, 171]. Homologs have been identified that are involved in $C_4$-dicarboxylate transport in *Wolinella succinogenes* [210], $C_4$-dicarboxylate-sensing in *Bacillus subtilis* [13], and glutamate transport in *Synechocystis* sp. [169]. It became evident that these systems are not only involved in uptake of organic anions, when the TeaABC system of *Halomonas elongata* was shown to transport the compatible solutes ectoine and hydroxyectoine [66].

The discovery of binding protein-dependent secondary transporters has raised questions concerning their transport mechanism as well as their evolutionary relationship to both the 'classical' secondary and the binding protein-dependent primary transporters (see 'Introduction'). These questions, in return, fuelled the research that is presented in this thesis.

**... in *Escherichia coli*: Quest for an elusive substrate**

The genome of the mesophilic Gram-negative laboratory bacterium *E. coli* K-12 contains one member of the family of binding protein-dependent secondary transporters, encoded by the *yiaMNO* genes [53, 59]. As this organism is more easily accessible with genetic and biochemical tools than the aforementioned *Rhodobacter* species, the YiaMNO transporter seemed a perfect system for answering the questions raised above. Thus, the three structural genes encoding the transporter were cloned, and preliminary expression studies were performed [Chapter 2]. However, investigation of the molecular mechanism of any transporter requires the use of a transported substrate 'X'. Unfortunately, identification of this compound 'X' turned out to be rather more difficult than was envisaged at the start.

At first, as found for the homologous transporters, it was thought that the substrate for the YiaMNO transporter might be an organic anion, possibly a $C_4$-dicarboxylate. Deletion mutants of the *yiaMNO* genes were constructed in various genetic backgrounds, but growth experiments performed on minimal media containing a broad range of varying carbon sources, and transport studies using radiolabeled compounds, did not point at a role of this system in $C_4$-dicarboxylate transport [48, Chapter 4]. In addition, binding studies making use of the purified periplasmic substrate-binding component YiaO and available radiolabeled compounds did not identify a substrate [Chapter 4]. A role in the accumulation of compatible solutes as described for the TeaABC system could also not be assigned to the YiaMNO transporter [Chapter 3]. Only recently, the identity of the elusive compound 'X' was unveiled.

The *yiaMNO* genes are part of the *yiaJKLMNOPQRS* gene cluster that has been implicated in carbohydrate metabolism [17, 183, Chapter 4 and 5]. Expression of this gene cluster is tightly regulated, and is not induced by any of the 80 potential substrates tested by Ibañez *et al*. [87]. Furthermore, unpublished binding studies performed with

purified YiaO suggested that the substrate might be a pentose [98]. Making use of a mutant *E. coli* K-12 strain that constitutively expresses the *yiaK-S* genes [17, 183], an elaborate metabolic and biochemical screen revealed that the rare pentose L-xylulose (L-*threo*-2-pentulose) is a substrate for the YiaMNO transporter [Chapter 4]. However, since L-xylulose does not induce expression of the system, the possibility cannot be excluded that that other substrates are recognized by the YiaMNO transporter as well.
Remarkably, although the 'wild type' strain MC4100 expresses the *yiaMNO* genes in a growth phase dependent manner [Chapter 3], this strain is unable to grow on L-xylulose when present as the sole carbon source [Chapter 4]. In this regard, a second L-xylulose uptake activity has been found in strains JA134 and its Δ*yiaMNO* derivative TP018. No firm conclusions about the relative contributions of either system can be made at this time. However, it is possible that the second system is not active in strain MC4100, and that in this strain the transport activity of YiaMNO alone does not support growth.

The identification of L-xylulose as a substrate of the YiaMNO transporter marked an important advance in the research on binding protein-dependent secondary transporters, as it demonstrated that apart from organic anions and compatible solutes these systems are able to take up carbohydrates. Unfortunately, it did not provide us with a useful tool for studying the molecular mechanism of the transporter as L-xylulose is not available in radioactive form, nor can it be be stably radiolabeled [Chapter 4, and unpublished data]. Thus, conventional transport studies using a purified system in proteoliposomes could not be performed.

**What is the fate of L-xylulose in *E. coli* K-12?**

L-xylulose can support growth as the sole carbon source only in strains that constitutively express the *yiaK-S* gene cluster [Chapter 4]. The genes located immediately downstream of the *yiaMNO* transporter genes encode enzymes with functions in carbohydrate metabolism [Chapter 5]. Thus, the proposed and proven functions of these enzymes allow some speculation on the fate of L-xylulose following internalisation. In addition to the *yiaP-S* genes, homologous enzymes from the *yjfQRSTUVWX* gene cluster of *E. coli* K-12 are taken into account here [Chapter 5]. The *yjfSTU* genes were recently shown to encode a phospho*enol*pyruvate:carbohydrate phosphotransferase system (PTS) specific for L-ascorbate [248], and the internalized

product L-ascorbate 6-phosphate is metabolized further by the enzymes encoded by the *yjfVWX* genes [244, 248]. These enzymes have been shown to have functions similar to the YiaQRS enzymes [Chapter 5] and are expressed in the strains used in our studies (T.H. Plantinga, unpublished results).

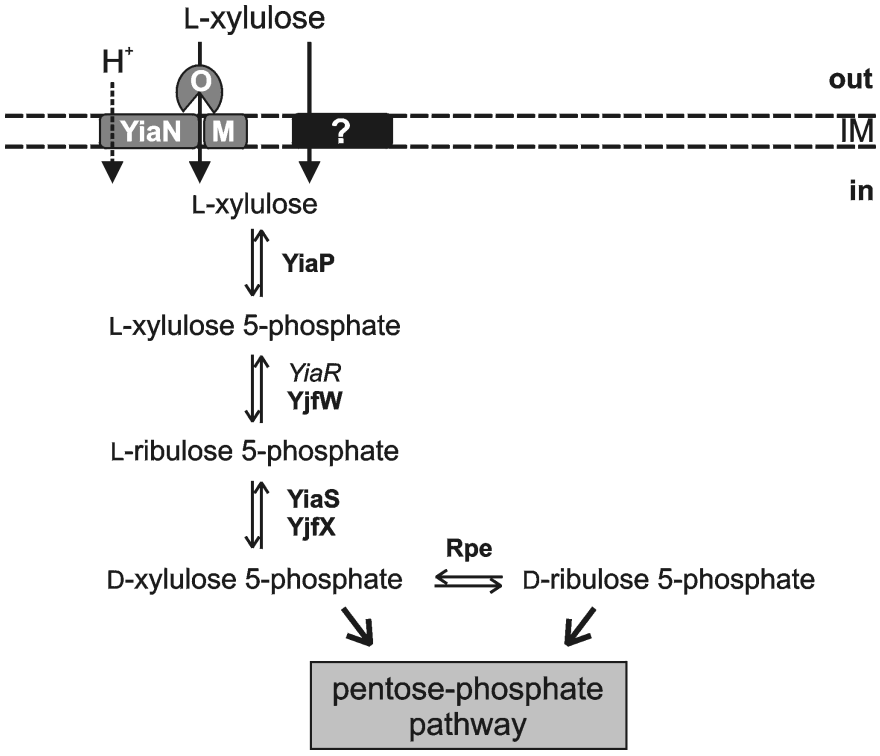



**Figure 1.** Putative degradation pathway for the substrate of the YiaMNO transporter (see text for details). L-xylulose is taken up via the YiaMNO transporter and/or a second as of yet unidentified system (question mark) [Chapter 4]. Bold: demonstrated function; Italic: proposed function; 'IM': inner membrane; 'out': periplasm; 'in': cytoplasm.

A possible degradative pathway for L-xylulose is proposed here (Fig. 1). L-xylulose enters the cell via the YiaMNO transporter, and/or a second unidentified system (Fig. 1) [Chapter 4] and is phosphorylated by YiaP [183]. The product L-xylulose 5-phosphate is first 3-epimerized to L-ribulose 5-phosphate (possibly by YiaR, although this activity has not been demonstrated, and/or YjfW), which is subsequently 4-epimerized to D-xylulose 5-phosphate (by YiaS and/or YjfX) [88, 244]. This compound then enters the non-oxidative branch of the pentose-phosphate pathway (PPP) [60]. Alternatively, the essential *E. coli* enzyme ribulose 5-phosphate epimerase Rpe [127] converts D-xylulose 5-phosphate into D-ribulose 5-phosphate, which subsequently enters the PPP (Fig. 1).

Additional branches of this pathway have been proposed by Reizer *et al*. [173]. In one proposed step, YiaS directly epimerizes L-xylulose 5-phosphate to D-ribulose 5-phosphate, which would then enter the PPP. However, YiaS was shown not to perform this epimerase activity [88]. Alternatively, it has been suggested that D-ribulose 5-phosphate is converted to D-arabino 3-hexulose 6-phosphate by YiaQ, and subsequently epimerized to D-fructose 6-phosphate by YiaR for entry into a glycolytic or oxidative pathway [173], yet this formaldehyde lyase-activity could not be demonstrated for YiaQ [243, and: J. Badia and J. Aguilar (personal communication)]. As these alternative metabolic routes have been discarded, they were left out of the model presented in Fig. 1 for reasons of clarity.

The true physiological function of the YiaMNO transporter remains unclear. L-xylulose uptake and metabolism by *E. coli* K-12 has not been studied before, and hardly any reports exist on the L-xylulose metabolism in other (micro)organisms. In humans, individuals suffering from essential pentosuria cannot reduce L-xylulose to xylitol and thus excrete large amounts of the pentose into urine. The condition has in the past been mistaken for diabetes, but appears to be harmless [112].

The phenotypical changes resulting from deleting the transporter cannot be easily explained by its L-xylulose transport funtion [Chapter 3]. These changes may result from alterations in the cell envelope, for example in the lipopolysaccharide (LPS). L-xylulose has been described as a building block of the LPS of the Gram-negative pathogen *Yersinia enterocolitica* [64] but not *E. coli* K-12. Interestingly, in the PPP D-xylulose 5-phosphate, a derivative of L-xylulose in the pathway described above (Fig. 1), together with D-ribose 5-phosphate is converted to D-sedoheptulose 7-phosphate by a transketolase (Fig. 2) [60]. The latter compound is a precursor of core LPS biosynthesis (Fig. 2, see figure legend for details) [102]. Interestingly, deletion of the GmhB enzyme results in reduction, but not a complete loss, of heptose in the LPS core, indicating the presence of an additional system in *E. coli* K-12 that partially compensates for this deletion. However, although a mutation that affects this pathway is not lethal for these cells, these mutants are more sensitive to challenges such as high bile salt concentrations [102]. Therefore, although at least nine enzymatic conversions separate L-xylulose from LPS core biosynthesis, it would be interesting to see whether or not the composition of the LPS is affected by the deletion of the YiaMNO transporter.

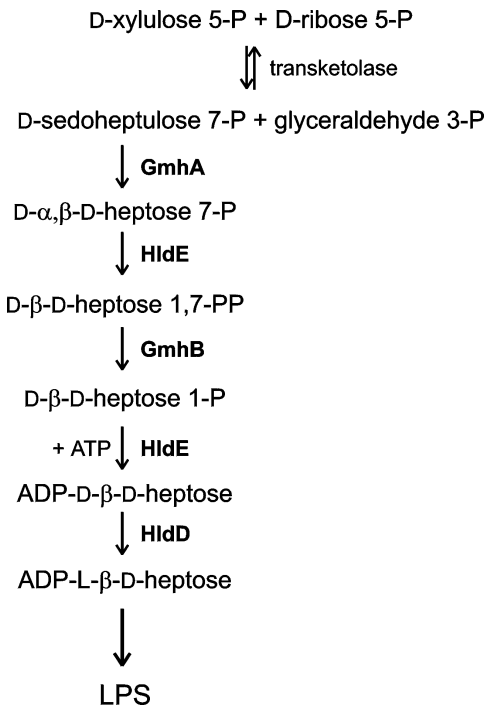



**Figure 2.** Proposed pathway for the biosynthesis of ADP-L-β-D-heptose, a LPS core precursor [102]. D-xylulose 5-phosphate and D-ribose 5-phosphate are converted into glyceraldehyde 3-phosphate and sedoheptulose 7-phosphate by a transketolase [60]. The latter compound is then converted to ADP-L-β-D-heptose via 5 enzymatic steps. GmhA: sedoheptulose 7-phosphate isomerase; HldE: bifunctional D-β-D-heptose 7-phosphate kinase/D-β-D-heptose 1-phosphate adenylyltransferase; GmhB: D-α,β-D-heptose 1,7-bisphosphate phosphatase; HldD: ADP-D-β-D-heptose epimerase [102]

## Perspectives

Due to the elusive nature of the transported substrate of the YiaMNO transporter, the questions that were raised at the start of this research have remained largely unanswered. At this time, the role of the pmf and/or the smf in binding protein-dependent transport by YiaMNO remains unclear. Both the kinetic characteristics and the molecular mechanism of transport could not be addressed. In addition, as the large membrane domain most likely is the permease, the function of the small inner membrane domain is obscure. One may speculate that the small subunit is involved in the interaction with the extracellular solute binding protein. These problems require a purified and functionally reconstituted system that can be used for detailed transport assays.

Another topic is the effect of the *yiaMNO* deletion on bacterial aggregation and biofilm formation [Chapter 3]. In this regard, it would be interesting to investigate not only the composition of the LPS, but also that of the exopolysaccharide (EPS) of both mutant and parent strains, as either one of these may have been affected by the deletion. Such experiments will possibly shed more light on the physiological function of L-xylulose uptake and metabolism by *E. coli* K-12. Moreover, expression of the system during these processes would have to be investigated. Whole genome analyses showed no significant change in expression under varying physiological circumstances including biofilm formation (see Chapter 3, 'Discussion'), but these studies made use of different strains and growth media, which all have an effect on results.

Finally, since the substrate-binding domain YiaO can be obtained in large quantities with relative ease, crystallization of this component for structural purposes could be an option. Although a large number of substrate-binding domains from members of the ABC superfamily has been crystallized, none of these proteins share significant sequence homology with members of the binding protein-dependent secondary transporters (see 'Introduction') [171]. It would be interesting to see whether these proteins are made up of a similar fold. In addition, a greater challenge would be to overexpress, purify and crystallize the YiaM-YiaN membrane complex, and see if a model for docking of the substrate-binding domain can be derived from such data (see 'Introduction'), as has been done for the Vitamin $B_{12}$-transporter Btu of *E. coli* [27].

**Concluding remarks**

To our knowledge, at least 136 putative substrates for the *yiaK-S* system have been tested, either as inducers of expression [87], or as transported substrates [Chapter 4]. Eventually, this 'quest' for the substrate has resulted in the identification of at least one compound that is transported by the YiaMNO transporter: L-xylulose. This now opens the path for further detailed studies on the structure and function of this intriguing class of transporters.

# *Chapter 7: Summary*

*"42."* [3]

Binding protein-dependent secondary, or tripartite ATP-independent periplasmic (TRAP), transporters combine distinguishing features from both the 'classical' secondary transporters and the binding protein-dependent primary transporters. The *Escherichia coli* K-12 genome contains one member of this recently discovered transport protein family. Since only a few members of this family have been functionally characterized to date, the *E. coli* K-12 homologue encoded by the *yiaMNO* genes became the focus of the research described in this thesis.

To enable studies on the kinetics and mechanism of transport, an overexpressed, purified and functionally reconstituted transporter is required. Chapter 2 describes our efforts to obtain such a system for the YiaMNO transporter. First, the *yiaMNO* genes were cloned and overexpressed in *E. coli*. Preliminary data show that the membrane proteins YiaM and YiaN could only be expressed from a synthetic operon that contained at least both the *yiaM* and the *yiaN* genes. This suggests that expression of the small membrane domain YiaM stabilized the expression of the large membrane domain YiaN, and/or vice versa. In addition, the substrate-binding protein YiaO could readily be overexpressed in the periplasm of *E. coli*, and was processed at the consensus Gram-negative signal sequence cleavage site. A C-terminal 6xHistidine-tag facilitated rapid purification of the protein from the periplasmic fraction.

Apart from a functionally reconstituted transporter, transport studies also require a substrate that can be easily monitored, preferably at small quantities. The majority of transport assays therefore make use of radioactively labeled derivatives of the transported substrate. As no such substrate is available for the YiaMNO transporter to date (see below), no further attempts were made to optimize the overexpression, purification and reconstitution procedure that was initiated in Chapter 2.

Chapter 3 describes our efforts to understand the physiological function of the YiaMNO transporter. *E. coli* K-12 strain MC4100 expresses the *yiaMNO* genes in a growth phase-dependent manner, and highest expression levels were found in cells that had been in stationary phase for 24h. Thus, the expression pattern suggests that the

system may play a role in scavenging of scarce substrate during the stationary growth phase. Strikingly, although a role of the system in carbohydrate uptake is described in this thesis (see below), expression is not repressed by glucose in the medium. In an attempt to identify the transported substrate of the YiaMNO transporter, an unmarked chromosomal deletion mutant of the *yiaMNO* genes was constructed in strain MC4100. This mutant was used in an elaborate series of growth experiments, testing over 100 potential substrates (see below, and Chapter 4). Although the function of the YiaMNO-mediated transport activity could not be inferred, the growth experiments show that deletion of the genes encoding the transporter results in some intriguing effects on the whole cell physiology of the organism. The deletion affected very diverse phenomena, such as the transition from exponential growth to the stationary phase, high salt survival, and biofilm formation by the mutant strain. Although the system clearly is expressed during stationary phase, no induction of a promoter-*lacZ* reporter construct was observed in the presence of high salt, or various other stressful conditions.
The phenotypic effects that were observed are rather diverse, and the underlying pathways are complex, but the observations made with the evolved *Ralstonia* sp. populations that have lost the *yiaMNO* genes (see Chapter 3) generally support our findings. Nevertheless, further studies are required to address the physiological role of the YiaMNO transporter.

The three genes encoding the transporter are located in the *yiaJKLMNOPQRS* gene cluster, which has been implicated in carbohydrate metabolism. Chapter 4 describes how *E. coli* strain JA134, that constitutively expresses the *yiaK-S* gene cluster, metabolized the rare pentose L-xylulose, whereas its parent strain ECL1 did not. Deletion of the *yiaMNO* transporter genes in strain JA134 reduced L-xylulose metabolism. Since L-xylulose cannot be radioactively labeled, indirect methods were used to measure substrate binding and transport. We found that the periplasmic substrate-binding protein YiaO is able to bind L-xylulose. In addition, this protein facilitates the uptake of L-xylulose by spheroplasts that express the membrane domains YiaM and YiaN. A second L-xylulose uptake activity is present in *E. coli* K-12, but was only detected in cells that constitutively express the genes located downstream of the transporter in the cluster.

The genomic survey described in Chapter 5 revealed that the clustering of the *yiaQRS* genes is conserved throughout a range of Gram-positive and Gram-negative bacteria, most of which are human pathogens. Strikingly, in each cluster there is a solute transport system present, although the class of transporter varies. In all cases, the genomic organization suggests that the substrate of the transporter is a carbohydrate that is phosphorylated either during transport, when there is a phospho*enol*pyruvate:carbohydrate phosphotransferase system (PTS) present, or immediately following transport by the kinase encoded within the cluster. Apart from the clustering of these genes the gene order is extremely conserved, suggesting that these enzymes specify similar functions.

The discovery of L-xylulose as a substrate of the YiaMNO transporter does not easily elucidate the physiological function of the transporter, as it does not explain the phenotype of the deletion mutant described in Chapter 3. In addition, since L-xylulose is unable to induce the expression of the YiaMNO transporter, while expression is not repressed by glucose, a major role of this system in carbon source uptake and utilization is not evident. However, it does significantly expand the range of potential substrates for binding protein-dependent secondary transporters from organic anions and compatible solutes to carbohydrates.

# *Hoofdstuk 8: Samenvatting*

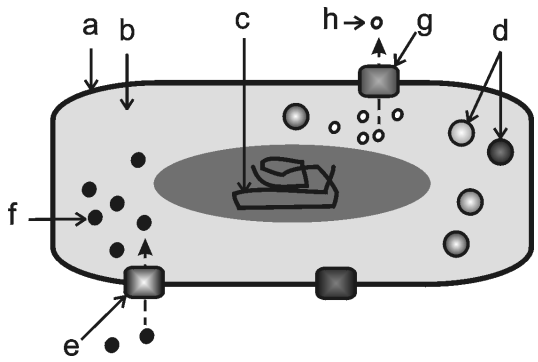
## Bindingseiwitafhankelijk secondair transport in *Escherichia coli*: Zoektocht naar een ongrijpbaar substraat

*"Omsk is een mooie stad, maar net iets te ver weg."* [55]

Het woord 'biologie' stamt uit het Grieks en betekent 'wetenschap van de levende materie'. Microbiologie is de wetenschap van de kleine levende materie (micro = klein), en onderzoekt alle processen die plaatsvinden in zogenaamde 'micro-organismen' zoals gisten en bacteriën. Het onderzoek dat is beschreven in dit proefschrift is uitgevoerd in een laboratorium waar wordt onderzocht hoe micro-organismen enerzijds voedingsstoffen opnemen uit hun omgeving, en anderzijds afvalstoffen weer uitscheiden. Dit transport vindt plaats over celmembranen die een natuurlijke barrier vormen die het binnenste van de cel scheidt van het buitenmedium, en wordt uitgevoerd door speciale eiwitten: de 'transporteiwitten' (zie paragraaf 2).
Dit proefschrift beschrijft het onderzoek naar de functie van een uniek, bindingseiwitafhankelijk secondair transporteiwit van de darmbacterie *Escherichia coli*.

### 1. Het model voor 'de' bacteriële cel: *Escherichia coli*

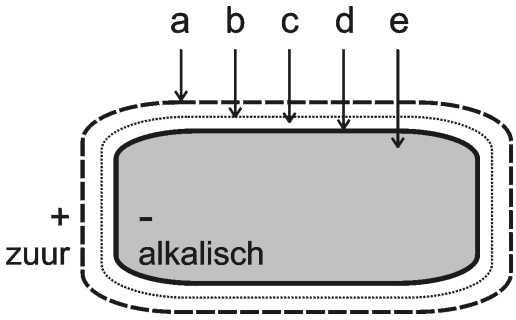
*Escherichia coli* K-12 is een in principe onschuldige darmbacterie die al decennia lang wereldwijd in laboratoria wordt gebruikt als modelorganisme voor microbiologisch onderzoek. Het genoom, dat wil zeggen: de complete verzameling van genen, van deze bacterie is in 1997 ontrafeld, waardoor onze kennis over dit organisme enorm is uitgebreid. Desondanks is zes jaar later nog steeds niet de functie bekend van alle genen die destijds zijn ontdekt. De bacterie ziet er onder de microscoop uit als een langwerpig, beweeglijk staafje, dat wordt voortgestuwd door zweepachtige uitsteeksels aan de buitenkant van de cel, de 'flagellen'. Figuur 1 laat een, op dit modelorganisme gebaseerd, sterk vereenvoudigd beeld zien van 'de' bacteriële cel. Deze is omgeven door de celenvelop (Fig. 1a, zie ook Fig. 2) die het binnenste van de cel, het cytoplasma (b), scheidt van de buitenwereld.




**Figuur 1.** Schematische weergave van 'de' bacteriële cel, gebaseerd op de modelbacterie *E. coli*. De celenvelop (a) (zie ook Fig. 2) schermt het cytoplasma (b) af van de buitenwereld. Hierin bevindt zich het erfelijke materiaal in de vorm van het DNA (c) dat in code de informatie bevat op basis waarvan eiwitten (d, e, g) worden gemaakt, met verschillende functies. Een aanzienlijk deel hiervan bevindt zich in de membraan en functioneert als transporteiwit voor de opname (e) of uitscheiding (g) van stoffen (f, h).

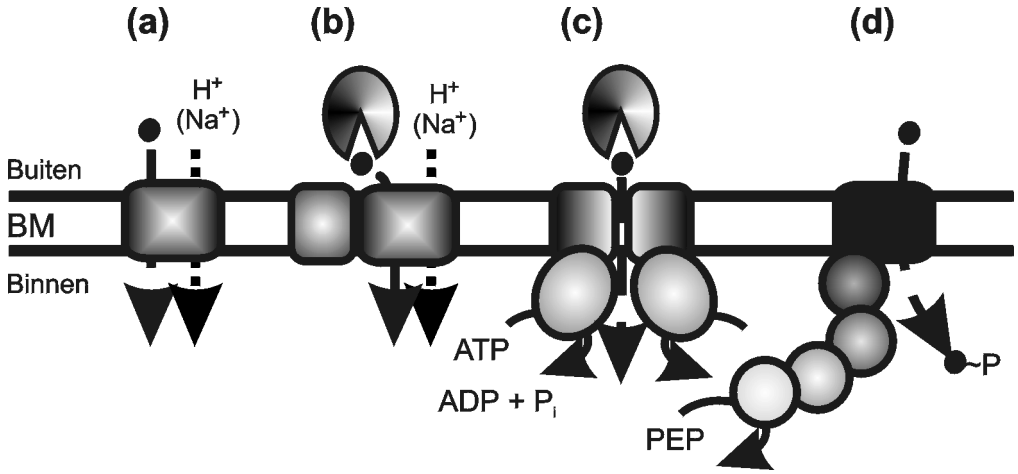
Het erfelijke materiaal van de cel, het desoxyribonucleïnezuur (DNA) (c), bevindt zich in het cytoplasma. De erfelijke informatie van de cel is in het DNA opgeslagen in de vorm van genen: korte stukken DNA die, in code, de informatie bevatten op basis waarvan eiwitten (d, e, g) worden gemaakt. Een groot aantal van deze eiwitten vervult een functie in de celmembraan (zie hieronder, en Fig. 2) als transporteiwit. Sommige zijn verantwoordelijk voor de opname (e) van bijvoorbeeld voedingsstoffen (f), terwijl andere weer gespecialiseerd zijn in de uitscheiding (g) van afvalstoffen (h). De stof die door een eiwit wordt getransporteerd wordt 'substraat' genoemd.

Figuur 2 geeft een schematisch overzicht van de membranen die een *E. coli* cel omgeven, de zogenaamde celenvelop, en de compartimenten die daardoor ontstaan. De buitenmembraan (a) bestaat grotendeels uit lipiden: vetachtige moleculen die een voor water ondoordringbare laag vormen. Er bevinden zich echter ook eiwitten in de buitenmembraan die hierin kanalen vormen, de 'porines', waardoor zich wel water kan verplaatsen met daarin opgeloste moleculen tot een bepaalde maximale grootte. Deze passeren ook gemakkelijk de celwand (b): een rigide webachtige structuur die verantwoordelijk is voor vormbehoud van de cel. De celwand bevindt zich in het periplasma (c), het compartiment tussen buiten- en binnenmembraan (d) wat een gel-achtig mengsel van water en daarin opgeloste stoffen bevat.




**Figuur 2.** Schematische weergave van de celenvelop van een *E. coli* K-12 cel. Er bevinden zich relatief meer protonen aan de buitenkant van de cel, hierdoor is het externe milieu 'zuur' ten opzichte van het 'alkalische' cytoplasma. Bovendien is het cytoplasma negatief geladen ten opzichte van het milieu. a) buitenmembraan; b) celwand; c) periplasma; d) binnenmembraan; e) cytoplasma.

Hierin bevinden zich bijvoorbeeld ook de losse, substraatbindende componenten van het type eiwit wat in dit proefschrift is onderzocht (zie paragraaf 2). De binnenmembraan (d) bestaat uit een laag lipiden met daarin vele eiwitten met verschillende functies, waaronder transporteiwitten. Doordat de binnenmembraan geen water met opgeloste stoffen doorlaat, en daarmee het cytoplasma (e) afscheidt van de buitenwereld, vervult deze een belangrijke functie. Als er zich namelijk aan de buitenkant van de cel meer moleculen van een bepaalde stof bevinden dan aan de binnenkant, ontstaat er een kracht die deze moleculen als het ware de cel in wil trekken, en vice versa. De binnenmembraan houdt gradiënten van moleculen in stand door ze niet door te laten. Dit geldt bijvoorbeeld ook voor kleine geladen moleculen, 'ionen', zoals protonen ($H^+$) en natrium ionen ($Na^+$). Door de samenstelling van het cytoplasma van *E. coli* heeft dit een negatieve lading ten opzichte van het externe milieu (Fig. 2). Bovendien bevinden zich relatief meer protonen aan de buitenkant van de cel dan in het cytoplasma, waardoor het externe milieu van de cel 'zuur' is ten opzichte van het 'alkalische' cytoplasma. De combinatie van de concentratiegradiënt van protonen en de ladingsverschillen resulteert in een kracht op de protonen, die deze als het ware het cytoplasma in drijft: de protonen drijvende kracht (pmf). Deze speelt een belangrijke rol bij de verschillende transportprocessen die plaatsvinden over de binnenmembraan (zie paragraaf 2).




**Figuur 3.** Vier verschillende klassen van bacteriële transporteiwitten (zie tekst voor details). a) Secondair transporteiwit. b) Bindingseiwitafhankelijk secondair transporteiwit. c) Bindingseiwitafhankelijk primair transporteiwit. d) Fosfo*enol*pyruvaat:suiker fosfotransferase systeem (PTS). '●': substraat; 'BM': binnenmembraan (in het geval van *E. coli* K-12 is 'buiten' het periplasma en 'binnen' het cytoplasma)

## 2. Transporteiwitten

Transporteiwitten kunnen worden verdeeld in verschillende klassen, op basis van zowel de bouw als het aandrijfmechanisme. De vier klassen van transporteiwitten die in bacteriën worden gevonden, zijn weergegeven in figuur 3. Secondaire transporteiwitten (a) bestaan uit één enkel eiwitmolecuul dat de gehele transportreactie uitvoert. Het transport wordt gedreven door de pmf, al speelt soms de gradiënt van natrium ionen ook een rol. Interessant is dat de richting van het transport afhankelijk is van de richting van de gradiënt. Als deze omgekeerd is kan het eiwit het substraat uitscheiden. In sommige gevallen gaat het daarbij om een efficiënte uitwisseling: dan is de opname van een voedingsstof gekoppeld aan de uitscheiding van een afbraakproduct ervan.

Twee klassen van transporteiwitten maken gebruik van een losse component aan de buitenkant van de cel die substraten ook bij lage hoeveelheden weet te detecteren, te binden en vervolgens bij het eigenlijke transporteiwit in de binnenmembraan af te leveren: een substraat-bindend eiwit (b, c). Van systemen met een dergelijk eiwit wordt gedacht dat ze uitsluitend in een richting, namelijk de cel in, werkzaam zijn. Het belangrijkste verschil tussen beide bindingseiwitafhankelijke systemen zit in de aandrijving: secondaire systemen worden gedreven door de pmf (b), terwijl de primaire

systemen worden gedreven door de energie afkomstig uit de omzetting van de energierijke verbinding ATP naar ADP en organisch fosfaat ($P_i$) (c). Tot voor kort werd gedacht dat alle bindingseiwitafhankelijke transportsystemen tot de primaire klasse (c) behoorden. De ontdekking van dergelijke systemen die door de pmf worden gedreven vormt de basis van het onderzoek in dit proefschrift (zie paragraaf 2).
De vierde klasse van transportsystemen zijn de zogenaamde fosfo*enol*pyruvaat:suiker fosfotransferase systemen (PTS) (d). Dit type transportsysteem neemt voornamelijk suikers op. Tijdens het transport wordt de fosfaat-groep van fosfo*enol*pyruvaat (PEP) via een aantal eiwitten overgedragen aan het opgenomen suiker.

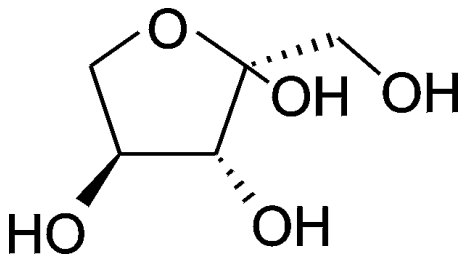
Ieder transporteiwit transporteert een (of meerdere) moleculen: het 'substraat'. Met behulp van dit substraat kan de activiteit en werking van het transporterende systeem worden onderzocht.

## 3. Het onderzoek in dit proefschrift

Het onderzoek dat is beschreven in dit proefschrift heeft tot doel inzicht te krijgen in de functie en de werking van het bindingseiwitafhankelijke secondaire transportsysteem van *E. coli*. Dit systeem wordt gevormd door drie eiwitten die worden gecodeerd door de *yiaMNO* genen. Het substraat van het YiaMNO systeem was bij het begin van dit onderzoek niet bekend, en bleek bijzonder moeilijk te identificeren (zie hieronder).

Hoofdstuk 2 van dit proefschrift beschrijft de klonering van de drie afzonderlijke eiwitten van de YiaMNO transporter. De experimenten met YiaM (het kleine membraaneiwit) en YiaN (het grote membraaneiwit) suggereren dat beide membraaneiwitten elkaar stabiliseren, en dus mogelijkerwijs in nauw contact met elkaar zijn in de membraan. Overproductie van YiaO bleek eenvoudig en het eiwit kon gemakkelijk geïsoleerd worden uit het periplasma van de producerende bacterie. In hoofdstuk 3 wordt de functie van het YiaMNO transporteiwit onderzocht. De productie van het eiwit door *E. coli* is het hoogst als de cellen de belangrijkste voedingsbron(nen) in hun medium hebben uitgeput en stoppen met delen (de 'stationaire fase' van de groei). Dit wordt niet negatief beïnvloed door de aanwezigheid van glucose, wat normaliter systemen remt die betrokken zijn bij de opname van een koolstofbron. De verwijdering van de *yiaMNO* genen heeft duidelijk invloed op de groei van de bacterie. Onder bepaalde kweekomstandigheden blijkt de mutant tot hogere celdichtheden te

groeien dan de uitgangsstam (het 'wild type'), maar het systeem waarmee de bacterie normaalgesproken deze populatiedichtheid signaleert is niet beïnvloed. De mutant groeit veel minder goed in aanwezigheid van een grote hoeveelheid zout. Bovendien is de mutant minder goed in staat om een 'biofilm' te vormen: hierbij hechten bacteriën zich aan een oppervlak en overwoekeren dit. Aangezien deze groeivorm in de natuur veel voorkomt en ook bij infecties een belangrijke rol speelt, is deze waarneming erg interessant. Tot op heden kunnen we de gevonden verschijnselen niet verklaren, maar er wel over speculeren (zie hieronder, en hoofdstuk 6).




**Figuur 4.** Schematische voorstelling van het substraat van de YiaMNO transporter: L-xylulose (L-*threo*-2-pentulose).

Hoofdstuk 4 vat de zoektocht naar het substraat van het YiaMNO transporteiwit samen. Meer dan 100 mogelijke substraten werden getest, maar uiteindelijk bleek informatie over het genoom van *E. coli* essentieel voor de identificatie van het substraat. De *yiaMNO* genen liggen namelijk in een groepje genen die eiwitten produceren die bepaalde suikers kunnen omzetten. Een daarvan, L-xylulose (Fig. 4), blijkt groeisubstraat te zijn voor gemuteerde cellen die deze eiwitten continu produceren. Mutanten van deze cellen die de YiaMNO eiwitten niet meer aanmaken groeien minder goed op L-xylulose. De cellen die zowel de transporter als de benodigde enzymen produceren consumeren L-xylulose en de mutant die het transporteiwit mist is hier minder goed in. Het bindingseiwit YiaO bindt L-xylulose specifiek en stimuleert de consumptie van dit suiker door 'spheroplasten' (cellen waarbij buitenmembraan en periplasma verwijderd zijn, maar de binnenmembraan nog wel in tact is) die YiaM en YiaN bevatten. Dit suggereert dat L-xylulose een substraat moet zijn van de YiaMNO transporter. Op dit moment is overigens niet uit te sluiten dat er nog meer substraten

zijn voor het eiwit. Bovendien worden de resultaten die zijn beschreven in hoofdstuk 3 niet onmiddellijk verklaard met de identiteit van dit substraat.

Hoofdstuk 5 beschrijft hoe de bovengenoemde enzymen die essentieel zijn voor het metabolisme van L-xylulose op dezelfde manier voorkomen in 26 andere bacteriën, waaronder een groot aantal potentiële ziekteverwekkers. Erg interessant is dat deze genen niet uitsluitend voorkomen in combinatie met een bindingseiwitafhankelijke secondaire transporter, maar met leden van elk van de vier families van bacteriële transportsystemen. Het lijkt er dus op dat deze systemen vergelijkbare functies hebben in opname en metabolisme van een suiker. De verschillen tussen deze genclusters beperkt zich in sommige gevallen tot enkele genen die als het ware lijken te zijn uitgewisseld. Dit is indicatie van de subtiliteit waarmee evolutionaire processen kennelijk kunnen verlopen.

In de hoofdstuk 6 worden bovenstaande hoofdstukken kort bediscussieerd. Er wordt een mogelijke afbraakroute voor L-xylulose beschreven, op basis van de (vermoedelijke) functies van de enzymen die vlak bij de *yiaMNO* genen op het genoom liggen. Bovendien wordt er ook gespeculeerd over een mogelijke rol van het suiker als bouwsteen voor een onderdeel van de buitenmembraan: de lipopolysaccharide. Een rol van L-xylulose aan het celoppervlak van de bacterie zou namelijk een mogelijke verklaring geven voor de effecten die beschreven zijn in hoofdstuk 3.

## 4. Tenslotte

Het onderzoek naar de functie van de bindingseiwitafhankelijke secondaire transporter van *E. coli* K-12 heeft geleid tot een onvoorziene, intensieve zoektocht naar het substraat ervan. Minstens 136 verbindingen zijn getest, zowel door ons als door andere onderzoeksgroepen, voordat bleek dat de YiaMNO transporter het zeldzame suiker L-xylulose transporteert. Deze ontdekking is bijzonder, omdat voorheen werd gedacht dat deze eiwitten geen suikers konden transporteren. Vragen naar het werkingsmechanisme van het systeem zijn onbeantwoord gebleven omdat L-xylulose zich niet leent voor gedetailleerde transportexperimenten. Bovendien blijft de functie van deze transportactiviteit onduidelijk. Zoals vaak het geval is in de wetenschap, heeft de beantwoording van een vraag dus vooral veel nieuwe vragen opgeworpen. De beantwoording van deze nieuwe vragen vormt een ware uitdaging.

# *Haadstik 9: Gearfetting*

## Biningsaaiwytôfhinklik sekondêr transport yn *Escherichia coli*: In syktocht nei in ûngrypber substraat

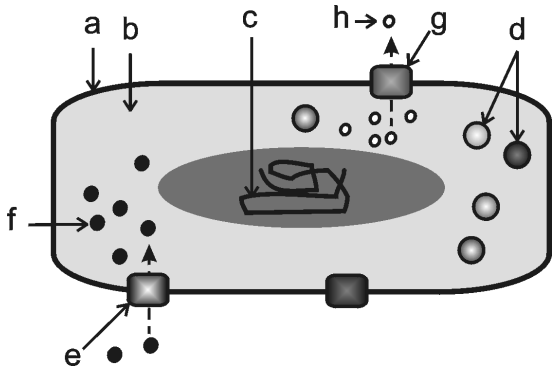
Titia H. Plantinga en Klaas Dantuma

*"As it net kin sa as't mat, dan mat it mar sa as't kin"* [1]

It Grykse wurd 'biology' betsjut 'wittenskip fan de libbene matearje', en 'mikrobiology' ('mikro' = 'lyts') is de wittenskip fan de lytse libbene matearje. Mikrobiologen ûndersykje alle prosessen dy't plakfine yn saneamde 'mikro-organismen', lykas gêsten en baktearjes. It ûndersyk dat is beskreaun yn dit proefskrift is útfierd yn in laboratoarium dêr't ûndersocht wurdt hoe't mikro-organismen oan de iene kant fiedingsstoffen opnimme út harren omjouwing, en oan de oare kant ôffalstoffen wer útskiede. Koart gearfette hâlde sy harren dwaande mei transport oer selmembranen (sjoch paragraaf 1). It transport wurdt útfierd toch spesiale aaiwiten: de 'transportaaiwiten' (sjoch paragraaf 2).
Dit proefskrift beskriuwt it ûndersyk nei de funksje fan in unyk, biningsaaiwytôfhinklik sekondêr transportaaiwyt fan de baktearje *Escherichia coli.*

### 1. It model foar 'de' bakteriële sel: *Escherichia coli*

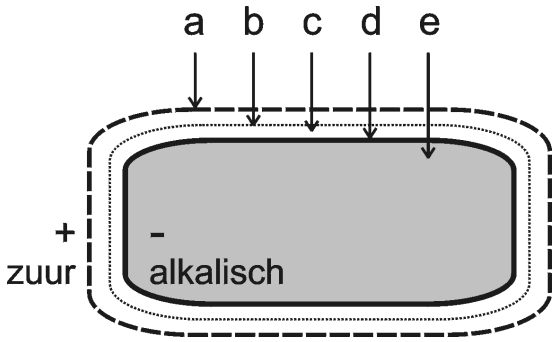
*Escherichia coli* is in yn prinsipe ûnskuldige termbaktearje dy't al desennia lang yn laboratoaria brûkt wurdt as modelorganisme foar mikrobilogysk ûndersyk. It 'genoom', dat wol sizze: de folsleine samling fan genen, fan dizze baktearje is yn 1997 útinoar raffele, en dêrmei is ús kennis oer dit organisme ûnbidich útwreide. Net tsjinsteande dat is 6 jier letter noch altyd net de funksje bekend fan alle genen dy't doedestiids ûntdekt binne.




**Figuer 1.** Skematyske werjefte fan 'de' bakteriële sel, basearre op de modelbaktearje *E. coli*. Sjoch tekst foar details.

De baktearje sjocht der ûnder de mikroskoop út as in langwerpich, beweeglik staafke, en figuer 1 jout hjir in sterk ferienfâldigd byld fan. De sel is omjûn troch de selslúf (Fig. 1a, sjoch ek Fig.2) dy't it binnenste fan de sel, it sytoplasma (b), skiedt fan de bûtenwrâld. It erflike materiaal fan de sel, it desoxyribonucleïnesoer (DNA) (c), sit yn it sytoplasma. De erflike ynformaasje fan de sel is yn it DNA opslein yn de foarm fan genen: koarte stikken DNA dy't yn koade de ynformaasje befetsje op grûn wêrfan de aaiwiten (d,e,g) makke wurde mei ûnderskate funksjes yn de sel. In grut oantal hjirfan is wurksum as transportaaiwyt (sjoch hjirûnder, en Fig. 2). Guon binne ferantwurdlik foar de opname (e) fan bygelyks fiedingsstoffen (f), wylst oaren spesialisearre binne yn de útskieding (g) fan bygelyks ôffalstoffen (h) (sjoch paragraaf 1.3). De stof dy't troch in aaiwyt transportearre wurdt, neame se 'substraat'.

Figuer 2 jout in skematysk oersjoch fan de membranen dy't in *E.coli* sel omjouwe: de selslúf. De bûtenmembraan (a) bestiet foar it grutste part út lipiden: fetachtige molekulen dy't in foar wetter ûntrochlitbere laach foarmje. Hjir befine har lykwols ek aaiwiten yn dy't kanalen foarmje, de 'porines', dêr't wetter mei dêryn oploste molekulen ta in bepaalde maksimale grutte harren wol ferpleatse kinne. Dit komt ek sûnder behindering lâns de selwand (b): in rigide webachtige struktuer dy't ferantwurdlik is foar foarmbehâld fan de sel. De selwand befynt him yn it 'periplasma' (c), de romte tusken bûten- en binnenmembraan (d) dy't in mingsel fan wetter en dêryn oploste stoffen befettet. Hjir sit bygelyks ek de losse, substraatbinende komponint yn fan it type aaiwyt dat yn dit proefskrift ûndersocht is (sjoch paragraaf 2).

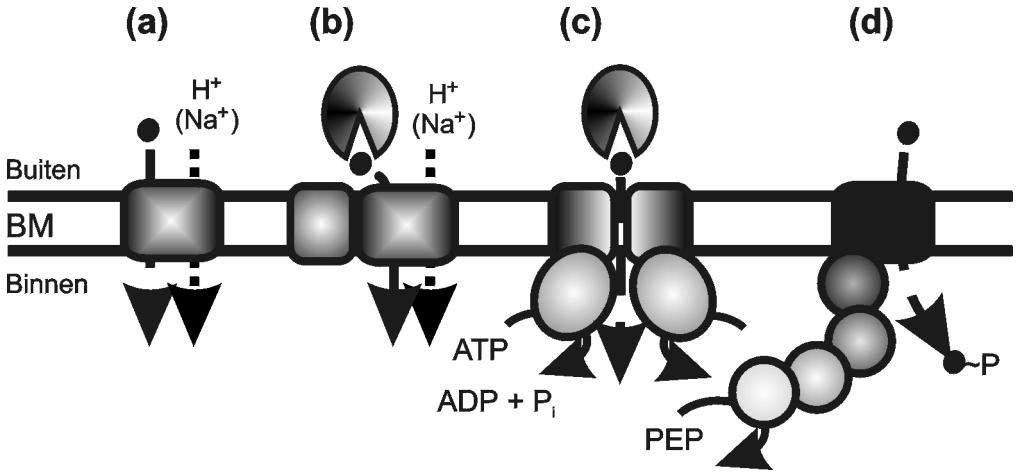



**Figuer 2**. Skematyske werjefte fan de selslúf fan in *E.coli* sel. Sjoch tekst foar details.

De binnenmembraan (d) bestiet út in laach lipiden met dêryn in hiel soad aaiwiten mei ûnderskate funksjes, lykas transportaaiwiten. Trochdat de binnenmembraan gjin wetter mei oploste stoffen trochlit ferfult dy in wichtige funksje as barriêre tusken bûtenwrâld en sytoplasma (e). As der him oan'e bûtenkant fan de sel mear molekulen fan in bepaalde stof befine as oan de binnenkant, ûntstiet der in krêft dy't dizze molekulen om samar te sizzen de sel ynlûke wol. De binnenmembraam hâldt dizze gradiïnt fan molekulen yn stân troch se net troch te litten. Dat jildt ek foar laden molekulen, 'ioanen', lykas protoanen ($H^+$) en natrium ioanen ($Na^+$). Troch de gearstalling fan it sytoplasma fan *E. coli* hat dit in negative lading oangeande de eksterne omjouwing (Fig.2). Boppedat binne der relatyf mear protoanen oan 'e bûtenkant fan de sel as yn it sytoplasma, sadat de eksterne omjouwing fan de sel relatief 'soer' is. De kombinaasje hjirfan resultearret yn in krêft op de protoanen dy't dizze om samar te sizzen it sytoplasma yn driuwt: de protoanendriuwende krêft (pmf). Dy spilet in wichtige rol by de ûnderskate transportprosessen dy't plakfine oer de binnenmembraam (sjoch paragraaf 2).

## 2. Transportaaiwiten

Transportaaiwiten binne yn mear of mindere mate spesialist: sy transportearje in bepaald (soarte) fan molekule, it saneamde substraat. Yn in oantal gefallen is in transportaaiwyt ek yn steat meardere molekulen dy't sterk op it favorite substraat lykje, te transportearjen.




**Figuer 3.** Fjouwer ûnderskate klassen fan bakteriële transportaaiwiten. a) Sekondêr transportaaiwyt. b) Biningsaaiwytôfhinklik sekondêr transportaaiwyt. c) Biningsaaiwytôfhinklik primêr transportaaiwyt. d) Fosfo*enol*pyruvaat:sûker fosfotransferase systeem (PTS). Sjoch tekst foar details. '**.**': substraat; 'BM': binnenmembraam (yn it gefal fan *E.coli* is 'bûten' it periplasma en 'binnen' it cytoplasma)

Transportaaiwiten wurde op basis fan opbou en oandriuwmechanisme ferdield yn ferskillende klassen. De fjouwer klassen dy't yn baktearjes fûn binne, binne werjûn yn figuer 3. Sekondêre transportaaiwiten (a) besteane út ien inkeld aaiwytmolekule dat de hiele transportreaksje útfiert. It transport wurdt dreaun troch de pmf en/of de gradiïnt fan natrium ioanen. Ynteressant is dat de rjochting fan it transport ôfhinklik is fan de rjochting fan de gradiïnt: wannear't dizze omkeard is kin it aaiwyt it substraat útskiede. Twa klassen fan transportaaiwiten meitsje gebrûk fan in losse komponint oan de bûtenkant fan de sel dy't it substraat detektearret, bynt en dan by it úteinlike transportaaiwyt yn de membraan ôfleveret; in biningsaaiwyt (b,c). Fan datsoarte fan transportaaiwiten wurdt tocht dat sy allinnich mar de sel yn transportearje. It wichtichtste ferskil tusken beide systemen sit yn de oandriuwing: sekondêre systemen meitsje gebrûk fan de pmf (b), wylst primêre systemen de enerzjy dy't ôfkomstich is út de omsetting fan de enerzjyrike ferbining ATP nei ADP en organysk fosfaat ($P_i$) brûke (c). Oant koartlyn waard tocht dat alle biningsaaiwytôfhinklike transportsystemen ta de primêre klasse (c) hearden, en de ûntdekking fan datsoarte systemen dy't troch de pmf dreaun wurde foarmet de basis fan dit ûndersyk yn dit proefskrift (sjoch paragraaf 3).

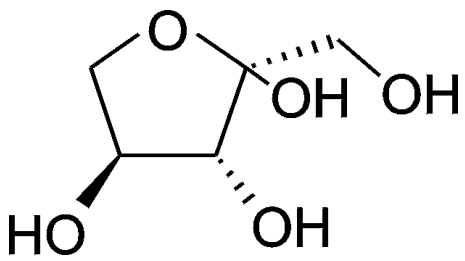
Dêrnei bestiet de fjirde klasse fan transportsystemen út de fosfo*enol*pyruvaat:sûker fosfotransferase systemen (PTS) (d). Dat type transportsysteem nimt benammen sûkers op. Under it transport wurdt de fosfaat-groep fan fosfo*enol*pyruvaat (PEP) fia in oantal aaiwiten oerdroegen oan it opnommen sûker.

## 3. It ûndersyk yn dit proefskrift

It ûndersyk dat is beskreaun yn dit proefskrift hat as doel ynsjoch te krijen yn de funksje en de wurking fan dit biningsaaiwytôfhinklike sekundêre transportsysteeem fan *E. coli*. De earste fraach yn dit ûndersyk hie te meitsjen mei de identiteit fan it substraat, omdat dêrmei de benoadigde eksperiminten útfierd wurde kinne. It modelorganisme fan de moderne mikrobiology, de beskikberens fan it komplete genoom fan dit organisme, en kennis fan de funksje fan twa besibbe transportaaiwiten út oare organismen soenen in optimale basis foarmje moatte foar dit ûndersyk. Dat wittenskiplik eksperimintearjen wat oars is as 'proefkes dwaan', die bliken doe't it oplossen fan dy earste fraach hast de hiele promoasjeperioade yn beslach naam.
It biningsaaiwytôfhinklike sekundêre transportsysteem fan *E.coli* wurdt foarme troch trije aaiwiten dy't kodearre wurde troch de *yiaMNO* genen. Haadstik 2 beskriuwt tariedende eksperiminten foar stúdzjes mei de *YiaMNO*-transporter. YiaM (it lytse membraamaaiwyt) en YiaN (it grutte membraamaaiwyt) kinne net los fan inoar produsearre wurde. Dit suggerearret dat beide membraamaaiwiten elkoar stabilisearje, en mooglik yn nau kontakt steane mei elkoar yn de binnenmembraan. It die bliken dat de produksje fan YiaO hiel ienfâldich wie, en dat it aaiwyt hiel maklik isolearre wurde koe út it periplasma. Dêr is gebrûk fan makke by de syktocht nei it substraat fan dit transportaaiwyt beskreaun yn haadstik 4.

Yn haadstik 3 wurdt de funksje fan it YiaMNO transportaaiwyt ûndersocht. Earst is sjoen oft, en sa ja op hokker momint, *E. coli* it aaiwyt produsearret. De produksje is op syn heechst as de sellen de wichtichste fiedingsboarne(n) yn harren medium útbrûkt hawwe en ophâlde mei dielen. De produksje fan de aaiwiten wurdt net negatyf beynfloede troch de oanwêzigens fan glukoaze, dat normaliter systemen ôfremmet dy't belutsen binnen by de opname fan in koalstofboarne. De ferwidering fan de *yiaMNO*-genen hat dúdlik ynfloed op de groei fan de baktearje. Under bepaalde

kweekomstannichheden docht bliken dat de mutant ta hegere seltichtheden groeit as de útgongsstam (it 'wild type'), mar it systeem dêr't de baktearje gewoanwei sprutsen dizze populaasjetichtens signalearret ('quorum sensing') is net beynfloede. In protte sâlt yn it medium remmet lykwols de groei fan de mutant. Boppedat is de mutant minder goed by steat om in 'biofilm' te foarmjen: yn dit gefal hechtsje baktearjes harren oan in opperflak en oerwaaksje dat. Sjoen it feit dat dizze groeifoarm yn'e natoer in soad foarkomt en ek by ynfeksjes in wichtige rol spilet, is dizze waarnimming tige ynteressant. Oant no ta kinne wy de fûne ferskynsels net ferklearje (sjoch hjirûnder, en haadstik 6).




**Figuer 4**. Skematyske foarstelling fan it substraat fan de YiaMNO transporter: L-xylulose (L-threo-2-pentulose).

Haadstik 4 fettet de syktocht nei it substraat fan it YiaMNO transportaaiwyt gear. Mear as 100 mooglike substraten waarden test. Eksperiminten mei mutearre baktearjes wiisden op it seldsume sûker L-xylulose (Fig. 4) as kandidaat-substraat. Wy hawwe bewize kinnen dat L-xylulose in substraat is foar de YiaMNO transporter. It biningsaaiwyt YiaO bynt L-xylulose spesifyk, en stimulearret de konsumpsje fan dit sûker troch 'spheroplasten' (sellen dêr't de bûtenmembraan en it periplasma fan ferwidere binne, mar de binnenmembraan noch wol fan yntakt is) dy't YiaM en YiaN befetsje. Op dit stuit is trouwens net út te sluten dat der noch mear substraten binne foar it aaiwyt. De funksje fan it systeem is nammentlik noch altyd ûndúdlik. Boppedat wurde de resultaten dy't beskreaun binne yn haadstik 3 net fuort-en-daliks ferklearrre mei de ydentiteit fan dit substraat.

Haadstik 5 beskriuwt hoe't de genen dy't deun by de *yiaMNO* genen lizze, en dy't wêzentlik binne foar it metabolisme fan L-xylulose, op krekt deselde wize foarkomme yn

26 oare baktearjes, dêr't in grut oantal fan yn wêzen sykteferwekkers is. Tige ynteressant is dat dizze genen net allinnich mar foarkomme yn kombinaasje mei in biningsaaiwytôfhinklike sekondêre transporter, mar mei leden fan alle fjouwer famyljes fan bakteriële transportsystemen (sjoch Fig. 3). It liket der dus op dat dizze systemen ferlykbere funksjes hawwe yn opname en metabolisme fan in sûker. De ferskillen tusken de genklusters beheine harren yn guon gefallen ta inkele genen dy't om samar te sizzen útwiksele blike te wêzen. Dat kin in yndikaasje wêze fan de subtiliteit dêr't evolusionêre prosessen blykber mei ferrinne kinne.

Yn haadstik 6 wurde boppesteande haadstikken koart bediskusiearre. Boppedat wurdt in mooglike ôfbrekrûte foar L-xylulose beskreaun, op grûn fan de (nei alle gedachten) funksjes fan de enzymen dy't deun by de *yiaMNO* genen op it genoom lizze. Der wurdt ek spekulearre oer in mooglike rol fan it sûker as boustien foar in ûnderdiel fan de bûtenmembraan: de lipopolysaccharide. In rol fan L-xylulose oan it selopperflak fan de baktearje soe nammentlik in mooglike ferklearring jaan foar de effekten dy't beskreaun binne yn haadstik 3.

## 4. As lêste

It ûndersyk nei de funksje fan de biningsaaiwytôfhinklike sekondêre transporter fan *E. coli* is útrûn op in ûnfoarsjoene, yntensive syktocht nei it substraat derfan. Op syn minst 136 ferbiningen binne úttest, sawol troch ús as troch oare ûndersyksgroepen, foardat bliken die dat de YiaMNO transporter it seldsume sûker L-xylulose transportearret. Dy ûntdekking hat it oantal mooglike substraten foar dit type aaiwyt drastysk útwreide, om't foarhinne tocht waard dat dy gjin sûkers transportearje koenen. Fragen nei it wurkingsmechanisme fan it systeem binne ûnbeäntwurde bleaun om't L-xylulose him net lient foar detaillearre transporteksperiminten.
Lykas wol faker it gefal is yn de wittenskip, hat de beäntwurding fan in fraach ús benammen nije fragen opsmiten. De beäntwurding fan dy nije fragen foarmet in wiere útdaging.

# *References*

*"Wow!"* [229]


1. Frisian proverb/Fries spreekwoord
2. Chinese proverb/Chinees spreekwoord

**A**

3. Adams, D. (1980) in: *The restaurant at the end of the universe*, Random House, New York
4. Albers, S.-V. (2001) Thesis, University of Groningen, Groningen, The Netherlands
5. Albers, S.-V., Elferink, M.G.L., Charlebois, R.L., Sensen, C.W., Driessen, A.J.M. & Konings, W.N. (1999) *J. Bacteriol.* **181**:4285-4291
6. Albers, S.-V., Szabó, Z. & Driessen, A.J.M. (2003) *J. Bacteriol.* **185**:3918-3925
7. Ames, G.F.-L. (1986) *Ann. Rev. Biochem.* **55**:397-425
8. Ames, G.F.-L. & Joshi, A.K. (1990) *J. Bacteriol.* **172**:4133-4137
9. Ames, G.F.-L., Liu, C.E., Joshi, A.K. & Nikaido, K. (1996) *J. Biol. Chem.* **271**:14264-14270
10. Antoine, R., Jacob-Dubuisson, F., Drobecq, H., Willery, E., Lesjean, S. & Locht, C. (2003) *J. Bacteriol.* **185**:1470-1474
11. Arfin, S.M., Long, A.D., Ito, E.T., Tolleri, L., Riehle, M.M., Paegle, E.S. & Hatfield, G.W. (2000) *J. Biol. Chem.* **275**:29672-29684
12. Arnold, C.N., McElhanon, J., Lee, A., Leonhart, R. & Siegele, D.A. (2001) *J. Bacteriol.* **183**:2178-2186
13. Asai, K., Baik, S.-H., Kasahara, Y., Moriya, S. & Ogasawara, N. (2000) *Microbiology* **146**:263-271
14. Ashwell, G. (1984) in: *Methods of enzymatic analysis. Volume VI. Metabolites I: Carbohydrates* (eds.: J. Bergmeier & M. Graßl), Verlag Chemie GmbH, Weinheim, pp. 477-483
15. Atlas, R.M. (1993) in: *Handbook of Microbiological Media* (ed.: L.C. Parks), Boca Raton pp. 533-533

**B**

16. Badia, J., Gimenez, R., Baldoma, L., Barnes, E., Fessner, W.-D. & Aguilar, J. (1991) *J. Bacteriol.* **173**:5144-5150
17. Badia, J., Ibañez, E., Baldoma, L. & Aguilar, J. (1998) *J. Biol. Chem.* **273**:8376-8381
18. Baneyx, F. & Georgiou, G. (1990) *J. Bacteriol.* **172**:491-494
19. Bassler, B.L. (2002) *Cell* **109**:421-424
20. Bayer, M.E. & Bayer, M.H. (1994) in: *Bacterial cell wall* (eds.: J.-M. Ghuysen & R. Hakenbeck) Elsevier Science Publishers, Amsterdam, pp. 447-464
21. Björkman, A.J. & Mowbray, S.L. (1998) *J. Mol. Biol.* **279**:651-664

22. Blattner, F.R., Plunkett III, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. & Shao, Y. (1997) *Science* **277**:1453-1462
23. Bloch, A. (1995) in: *Academic Murphology*, Bruna bv, Utrecht
24. Bochner, B.R. & Savageau, M.A. (1977) *Appl. Env. Microbiol.* **33**:434-444
25. Bolhuis, H., Van Veen, H.W., Molenaar, D., Poolman, B., Driessen, A.J.M. & Konings, W.N. (1996) *EMBO J.* **15**:4239-4245
26. Boos, W. & Lucht, J.M. (1996) in: Escherichia coli *and* Salmonella*: cellular and molecular biology*, 2nd edition (eds.: F.C. Neidhardt, R.I. Curtiss, J.L. Ingraham, C.C. Lin, K.B. Low, B. Magasnik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger) ASM Press, Washington D.C., pp. 1175-1209
27. Borths, E.L., Locher, K.P., Lee, A.T. & Rees, D.C. (2002) *Proc. Natl. Acad. Sci. USA* **99**:16642-16647
28. Bott, M., Meyer, M. & Dimroth, P. (1995) *Mol. Microbiol.* **18**:533-546
29. Brass, J.M., Higgins, C.F., Foley, M., Rugman, P.A., Birmingham, J. & Garland, P.B. (1986) *J. Bacteriol.* **165**:787-794

**C**

30. Cadieux, N., Bradbeer, C., Reeger-Schneider, E., Köster, W., Mohanty, A.K., Wiener, M.C. & Kadner, R.J. (2002) *J. Bacteriol.* **184**:706-717
31. Campos, E., Aguilar, J., Baldoma, L. & Badia, J. (2002) *J. Bacteriol.* **184**:6065-6068
32. Cao, G., Cheng, S., Whitley, P., von Heijne, G., Kuhn, A. & Dalbey, R.E. (1994) *J. Biol. Chem.* **269**:26898-26903
33. Chami, M., Steinfels, E., Orelle, C., Jault, J.-M., Di Pietro, A., Rigaud, J.-L. & Marco, S. (2002) *J. Mol. Biol.* **315**:1075-1085
34. Chang, G. & Roth, C.B. (2002) *Science* **293**:1793-1800
35. Chaudhuri, B.N., Ko, J., Park, C., Jones, T.A. & Mowbray, S.L. (1999) *J. Mol. Biol.* **286**:1519-1531
36. Chen, J., Sharma, S., Quiocho, F.A. & Davidson, A.L. (2001) *Proc. Natl. Acad. Sci. USA* **98**:1525-1530
37. Chen, X., Schauder, S., Potier, N., Van Dorsselaer, A., Pelczer, I., Bassler, B.L. & Hughson, F.M. (2002) *Nature* **415**:545-549
38. Conan Doyle, A. (1890) in: *The sign of four*, The Strand Magazine, London

**D**

39. Danese, P.N., Pratt, L.A. & Kolter, R. (2000) *J. Bacteriol.* **182**:3593-3596
40. Dassa, E. & Bouige, P. (2001) *Res. Microbiol.* **152**:211-229
41. Dassa, E. & Hofnung, M. (1985) *EMBO J.* **4**:2287-2293
42. Davey, M.E., Caiazza, N.C. & O'Toole, G. (2003) *J. Bacteriol.* **185**:1027-1036

43. Davidson, A.L. (2002) *Science* **296**:1038-1040
44. Davidson, A.L., Laghaeian, S.S. & Mannering, D.E. (1996) *J. Biol. Chem.* **271**:4858-4863
45. Davidson, A.L., Shuman, H.A. & Nikaido, H. (1992) *Proc. Natl. Acad. Sci. USA* **89**:2360-2364
46. Davidson, A.L. (2002) *J. Bacteriol.* **184**:1225-1233
47. Davidson, A.L. & Nikaido, K. (1990) *J. Biol. Chem.* **265**:4254-4260
48. Davies, S.J., Golby, P., Omrani, D., Broad, S.A., Harrington, V.L., Guest, J.R., Kelly, D.J. & Andrews, S.C. (1999) *J. Bacteriol.* **181**:5624-5635
49. Detmers, F.J.M. (2001) in: Thesis, University of Groningen, Groningen, The Netherlands
50. Diederichs, K., Diez, J., Greller, G., Müller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W. & Welte, W. (2000) *EMBO J.* **19**:5951-5961
51. Diez, J., Diederichs, K., Greller, G., Horlacher, R., Boos, W. & Welte, W. (2001) *J. Mol. Biol.* **305**:905-915
52. Dorel, C., Vidal, O., Prigent-Combaret, C., Vallet, I. & Lejeune, P. (1999) *FEMS Microbiol. Lett.* **178**:169-175
53. Driessen, A.J.M., Jacobs, M.H.J. & Konings, W.N. (1997) *Mol. Microbiol.* **24**:879-883
54. Driessen, A.J.M., Rosen, B.P. & Konings, W.N. (2000) *Trends Biochem. Sci.* **25**:397-401
55. Drs.P. (1974) in: *Dodenrit*, Polydor bv

**E**

56. Elliott, T. (1992) *J. Bacteriol.* **174**:245-253

**F**

57. Fath, M.J. & Kolter, R. (1993) *Microbiol. Rev.* **57**:995-1017
58. Fetsch, E.E. & Davidson, A.L. (2002) *Proc. Natl. Acad. Sci. USA* **99**:9685-9690
59. Forward, J.A., Behrendt, M.C., Wyborn, N.R., Cross, R. & Kelly, D.J. (1997) *J. Bacteriol.* **179**:5482-5493
60. Fraenkel, D.G. (1996) in: Escherichia coli *and* Salmonella*: cellular and molecular biology*, 2nd edition (eds.: F.C. Neidhardt, R.I. Curtiss, J.L. Ingraham, C.C. Lin, K.B. Low, B. Magasnik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger) ASM Press, Washington D.C., pp. 189-198
61. Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A. & Perret, X. (1997) *Nature* **387**:394-401
62. Fukuma-Kobayashi, K., Tateno, Y. & Nishikawa, K. (1999) *J. Mol. Biol.* **286**:279-290

**G**

63. Gorbulev, S., Abele, R. & Tampé, R. (2001) *Proc. Natl. Acad. Sci. USA* **98**:3732-3737
64. Gorshkova, R.P., Isakov, V.V., Kalmykova, E.N. & Ovodov, Y.S. (1995) *Carbohydrate Res.* **268**:249-255

65. Goswitz, V.C. & Brooker, R.J. (1995) *Protein Sci.* **4**:534-537
66. Grammann, K., Volke, A. & Kunte, H.J. (2002) *J. Bacteriol.* **184**:3078-3085
67. Grimard, V., Vigano, C., Margolles, A., Wattiez, R., Van Veen, H.W., Konings, W.N., Ruysschaert, J.-M. & Goormaghtigh, E. (2001) *Biochemistry* **40**:11876-11886
68. Gunn FJ, Tate, C.G. & Henderson, P.J.F. (1994) *Mol. Microbiol.* **12**:799-809

**H**

69. Hacksell, I., Rigaud, J.-L., Purhonen, P., Pourcher, T., Hebert, H. & Leblanc, G. (2002) *EMBO J.* **21**:3569-3574
70. Hall, J.A., Ganesan, A.K., Chen, J. & Nikaido, K. (1997) *J. Biol. Chem.* **272**:17615-17622
71. Hall, J.A., Gehring, K. & Nikaido, K. (1997) *J. Biol. Chem.* **272**:17605-17609
72. Hall, J.A., Thorgeirsson, T.E., Liu, J., Shin, Y.-K. & Nikaido, K. (1997) *J. Biol. Chem.* **272**:17610-17614
73. Hamblin, M.J., Shaw, J.G., Curson, J.P. & Kelly, D.J. (1990) *Mol. Microbiol.* **4**:1567-1574
74. Henderson, P.J.F. (1993) *Curr. Opin. Cell Biol.* **5**:708-721
75. Heyne, R.I., De Vrij, W., Crielaard, W. & Konings, W.N. (1991) *J. Bacteriol.* **173**:791-800
76. Hickman, J. & Ashwell, G. (1959) *J. Biol. Chem.* **234**:758-761
77. Higgins, C.F. (1992) *Annu. Rev. Cell. Biol.* **8**:67-113
78. Higgins, C.F. (2001) *Res. Microbiol.* **152**:205-210
79. Higgins, C.F., Hiles, I.D., Salmond, G.P., Gill, D.R., Downie, J.A., Evans, I.J., Holland, I.B., Gray, L., Buckel, S.D., Bell, A.W. & Hermodson, M.A. (1986) *Nature* **323**:448-450
80. Hirai, T., Heymann, J.A.W., Maloney, P.C. & Subramaniam, S. (2003) *J. Bacteriol.* **185**:1712-1718
81. Hirai, T., Heymann, J.A.W., Shi, D., Sarker, R., Maloney, P.C. & Subramaniam, S. (2002) *Nature Struct. Biol.* **9**:597-600
82. Hobot, J.A., Carlemalm, E., Villiger, W. & Kellenberger, E. (1985) *J. Bacteriol.* **160**:143-152
83. Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P. & Tainer, J.A. (2000) *Cell* **101**:789-800
84. Hsiao, C.-D., Sun, Y.-J., Rose, J. & Wang, B.-C. (1996) *J. Mol. Biol.* **262**:225-242
85. Hung, L.-W., Wang, L.X., Nikaido, H., Liu, P.-Q., Ames, G.F.-L. & Kim, S.H. (1998) *Nature* **396**:703-707
86. Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E. & Higgins, C.F. (1990) *Nature* **346**:362-365

**I**

87. Ibañez, E., Campos, E., Baldoma, L., Aguilar, J. & Badia, J. (2000) *J.Bacteriol.* **182**:4617-4624
88. Ibañez, E., Gimenez, R., Pedraza, T., Baldoma, L., Aguilar, J. & Badia, J. (2000) *J.Bacteriol.* **182**:4625-4627

**J**


89. Jackson, D.W., Simecka, J.W. & Romeo, T. (2002) *J. Bacteriol.* **184**:3406-3410
90. Jackson, D.W., Suzuki, K., Oakford, L., Simecka, J.W., Hart, M.E. & Romeo, T. (2002) *J. Bacteriol.* **184**:290-301
91. Jacobs, M.H.J., van der Heide, T., Driessen, A.J.M. & Konings, W.N. (1996) *Proc. Natl. Acad. Sci. USA* **93**:12786-12790
92. Jin, J., Guffanti, A.A., Beck, C. & Krulwich, T.A. (2001) *J. Bacteriol.* **183**:2667-2671
93. Jones, P.M. & George, A.M. (1999) *FEMS Microbiol. Lett.* **179**:187-202


**K**


94. Kaback, H.R. (1997) *Proc. Natl. Acad. Sci. USA* **94**:5539-5543
95. Kapatral, V., Anderson, I., Ivanova, N., Reznik, G., Los, T., Lykidis, A., Bhattacharayya, A., Bartman, A., Gardner, W., Grechkin, G., Zhu, L., Vasieva, O., Chu, L., Kogan, Y., Chaga, O., Goltsman, E., Bernal, A., Larsen, N., D'Souza, M., Walunas, T., Pusch, G., Haselkorn, R., Fontstein, M., Kyrpides, N. & Overbeek, R. (2002) *J. Bacteriol.* **184**:2005-2018
96. Karpowich, N., Martsinkevich, O., Millen, L., Yuan, Y.-R., Dai, P.L., MacVey, K., Thomas, P.J. & Hunt, J.F. (2001) *Structure* **9**:571-586
97. Karpowich, N.K., Huang, H.H., Smith, P.C. & Hunt, J.F. (2003) *J. Biol. Chem.* **278**:8429-8434
98. Kelly, D.J. & Thomas, G.H. (2001) *FEMS Microbiol. Rev.* **25**:405-424
99. Kennedy, K.A. & Traxler, B. (1999) *J. Biol. Chem.* **274**:6259-6264
100. Kihara, A., Akiyama, Y. & Ito, K. (1995) *Proc. Natl. Acad. Sci. USA* **92**:4532-4536
101. Kjærgaard, K., Schembri, M.A., Hasman, H. & Klemm, P. (2000) *J. Bacteriol.* **182**:4789-4796
102. Kneidinger, B., Marolda, C., Graninger, M., Zamyatina, A., McArthur, F., Kosma, P., Valvano, M.A. & Messner, P. (2002) *J. Bacteriol.* **184**:363-369
103. Knol, J., Veenhoff, L., Liang, W.-J., Henderson, P.J.F., Leblanc, G. & Poolman, B. (1996) *J. Biol. Chem.* **271**:15358-15366
104. Kokoeva, M.V., Storch, K.-F., Klein, C. & Oesterhelt, D. (2002) *EMBO J.* **21**:2312-2322
105. Kolter, R. (1999) *J. Bacteriol.* **181**:697-699
106. Koning, S.M., Konings, W.N. & Driessen, A.J.M. (2002) *Archaea* **1**:19-25
107. Koonin, E.V., Makarova, K.S. & Aravind, L. (2001) *Annu. Rev. Microbiol.* **55**:709-742
108. Korona, R., Nakatsu, C.H., Forney, L.J. & Lenski, R.E. (1994) *Proc. Natl. Acad. Sci. USA* **91**:9037-9041
109. Kreimer, D.I., Malak, H., Lakowicz, J.R., Trakhanov, S., Villar, E. & Shnyrov, V.L. (2000) *Eur. J. Biochem.* **267**:4242-4252
110. Kuan, G., Dassa, E., Saurin, W., Hofnung, M. & Saier, M.H.J. (1995) *Res. Microbiol.* **146**:271-278

**L**


111. Landmesser, H., Stein, A., Blüschke, B., Brinkmann, M., Hunke, S. & Schneider, E. (2002) *Biochim. Biophys. Acta* **1565**:64-72
112. Lane, A.B. (1985) *Biochem. Genet.* **23**:61-72
113. Lankat-Buttgereit, B. & Tampé, R. (2002) *Physiol. Rev.* **82**:187-204
114. Leenhouts, K., Buist, G., Bolhuis, A., ten Berge, A., Kiel, J., Mierau, I., Dabrowska, M., Venema, G. & Kok, J. (1996) *Molec. Gen. Genet.* **253**:217-224
115. Lerouge, I. & Vanderleyden, J. (2001) *FEMS Microbiol. Rev.* **26**:17-47
116. Lin, E.C.C. (1976) *Annu. Rev. Microbiol.* **30**:535-578
117. Linton, K.J. & Higgins, C.F. (1998) *Mol. Microbiol.* **28**:5-13
118. Liu, C.E. & Ames, G.F.-L. (1997) *J. Biol. Chem.* **272**:859-866
119. Liu, C.E., Liu, P.-Q. & Ames, G.F.-L. (1997) *J. Biol. Chem.* **272**:21883-21891
120. Liu, C.E., Liu, P.-Q., Wolf, A., Lin, E. & Ames, G.F.-L. (1999) *J. Biol. Chem.* **274**:739-747
121. Liu, P.-Q. & Ames, G.F.-L. (1998) *Proc. Natl. Acad. Sci. USA* **95**:3495-3500
122. Locher, K.P., Lee, A.T. & Rees, D.C. (2002) *Science* **296**:1091-1098
123. Lolkema, J.S., Enequist, H.G. & Van der Rest, M.E. (1994) *Eur. J. Biochem.* **220**:469-475
124. Lolkema, J.S., Poolman, B. & Konings, W.N. (1998) *Curr. Opin. Microbiol.* **1**:248-253
125. Loo, T.W., Bartlett, M.C. & Clarke, D.M. (2002) *J. Biol. Chem.* **277**:41303-41306
126. Loo, T.W. & Clarke, D.M. (2002) *Proc. Natl. Acad. Sci. USA* **99**:3511-3516
127. Lyngstadaas, A., Sprenger, G.A. & Boye, E. (1998) *Biochim. Biophys. Acta* **1381**:319-330


**M**


128. Magnusson, U., Chaudhuri, B.N., Ko, J., Park, C., Jones, T.A. & Mowbray, S.L. (2002) *J. Biol. Chem.* **277**:14077-14084
129. Maloney, P.C. & Wilson, T.H. (1996) in: Escherichia coli *and* Salmonella*: cellular and molecular biology*, 2nd edition (eds.: F.C. Neidhardt, R.I. Curtiss, J.L. Ingraham, C.C. Lin, K.B. Low, B. Magasnik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger) ASM Press, Washington D.C., pp. 1130-1148
130. Manting, E.H. & Driessen, A.J.M. (2000) *Mol. Microbiol.* **37**:226-238
131. Mao, B., Pear, M.R., McCammon, J.A. & Quiocho, F.A. (1982) *J. Biol. Chem.* **257**:1131-1133
132. Marger, M.D. & Saier, M.H. (1993) *Trends Biochem. Sci.* **18**:13-20
133. Margolles, A., Putman, M., Van Veen, H.W. & Konings, W.N. (1999) *Biochemistry* **38**:16298-16306
134. Marques, S. & Manzanera, M. (1999) *Mol. Microbiol.* **31**:1105-1113
135. Martins, L.O. & Santos, H. (1995) *Appl. Env. Microbiol.* **61**:3299-3303
136. Marty-Teysset, C., Posthuma, C., Lolkema, J.S., Schmitt, P., Divies, C. & Konings, W.N. (1996) *J. Bacteriol.* **178**:2178-2185
137. Mason, C.A. & Egli, T. (1993) in: *Starvation in bacteria* (ed.: S. Kjelleberg) Plenum Press, New York, pp. 81-102

138. Matsuyama, S., Akimaru, J. & Mizushima, S. (1990) *FEBS Lett.* **269**:96-100
139. Meijer, W.G., Arnberg, A.C., Enequist, H.G., Terpstra, P., Lidstrom, M.E. & Dijkhuizen, L. (1991) *Molec. Gen. Genet.* **225**:320-330
140. Merino, G. & Shuman, H.A. (1997) *J. Bacteriol.* **179**:7687-7694
141. Miller, J.H. (1972) in: *Experiments in molecular genetics*, Cold Spring Harbor Laboratory, New York
142. Miller, J.H. (1992) in: *A short course in bacterial genetics*, Cold Spring Harbor Laboratory, New York
143. Moody, J.E., Millen, L., Binns, D., Hunt, J.F. & Thomas, P.J. (2002) *J. Biol. Chem.* **277**:21111-21114
144. Mourez, M., Hofnung, M. & assa, E. (1997) *EMBO J.* **16**:3066-3077
145. Muiry, J.A.R., Gunn, T.C., McDonald, T.P., Bradley, S.A., Tate, C.G. & Henderson, P.J.F. (1993) *Biochem. J.* **290**:833-842

**N**

146. Nakatsu, C.H., Korona, R., Lenski, R.E., De Bruijn, F.J., Marsh, T.L. & Forney, L.J. (1998) *J. Bacteriol.* **180**:4325-4331
147. Nanavati, D., Noll, K.M. & Romano, A.H. (2002) *Microbiology* **148**:3531-3537
148. Neu, H.C. & Heppel, L.A. (1965) *J. Biol. Chem.* **240**:3685-3692
149. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997) *Prot. Engin.* **10**:1-6
150. Nikaido, K., Liu, P.-Q. & Ames, G.F.-L. (1997) *J. Biol. Chem.* **272**:27745-27752

**O**

151. O'Toole, G.A., Pratt, L.A., Watnick, P.A., Newman, D.K., Weaver, V.B. & Kolter, R. (1999) in: *Methods in enzymology* (eds.: R.A. Doyle, J.N. Abelson, and M.I. Simon) Academic Press, pp. 91-109
152. O'Toole, G., Kaplan, H.B. & Kolter, R. (2000) *Annu. Rev. Microbiol.* **54**:49-79
153. Obis, D., Guillot, A., Gripon, J.-C., Renault, P., Bolotin, A. & Mistou, M.-Y. (1999) *J. Bacteriol.* **181**:6283-6246
154. Oh, B.-H., Kang, C.-K., De Bondt, H., Kim, S.-H., Nikaido, K., Joshi, A.K. & Ames, G.F.-L. (1994) *J. Biol. Chem.* **269**:4135-4143
155. Otto, K. & Silhavy, T.J. (2002) *Proc. Natl. Acad. Sci. USA* **99**:2287-2292

**P**

156. Palin, M., Jones, T., Idle, E., Gilliam, T., Cleese, J. & Chapman, G. (1974) in: *Monty Python and the Holy Grail*
157. Pao, S.S., Paulsen, I.T. & Saier, M.H. (1998) *Microbiol. Mol. Biol. Rev.* **62**:1-34
158. Park, Y., Cho, Y.-J., Ahn, T. & Park, C. (1999) *EMBO J.* **18**:4149-4156
159. Paulsen, I.T., Nguyen, L., Sliwinski, M.K., Rabus, R. & Saier, M.H. (2000) *J. Mol. Biol.* **301**:75-100

160. Plantinga, T.H. (2003) Thesis, University of Groningen, Groningen, The Netherlands
161. Poelarends, G.J. & Konings, W.N. (2002) *J. Biol. Chem.* **277**:42891-42898
162. Pomposiello, P.J., Bennik, M.H.J. & Demple, B. (2001) *J. Bacteriol.* **183**:3890-3902
163. Poolman, B. & Konings, W.N. (1993) *Biochim. Biophys. Acta* **1183**:5-39
164. Postma, P.W., Lengeler, J.W. & Jacobson, G.R. (1993) *Microbiol. Rev.* **57**:543-594
165. Postma, P.W., Lengeler, J.W. & Jacobson, G.R. (1996) in: Escherichia coli *and* Salmonella*: cellular and molecular biology*, 2nd edition (eds.: F.C. Neidhardt, R.I. Curtiss, J.L. Ingraham, C.C. Lin, K.B. Low, B. Magasnik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger) ASM Press, Washington D.C., pp. 1149-1174
166. Prigent-Combaret, C., Brombacher, E., Vidal, O., Ambert, A., Lejeune, P., Landini, P. & Dorel, C. (2001) *J. Bacteriol.* **183**:7213-7223
167. Prigent-Combaret, C., Vidal, O., Dorel, C. & Lejeune, P. (1999) *J. Bacteriol.* **181**:5993-6002
168. Putman, M., Van Veen, H.W., Poolman, B. & Konings, W.N. (1999) *Biochemistry* **38**:1002-1008

**Q**

169. Quintero, M.J., Montesinos, M.L., Herrero, A. & Flores, E. (2001) *Genome Res.* **11**:2034-2040
170. Quiocho, F.A. & Ledvina, P.S. (1996) *Mol. Microbiol.* **20**:17-25

**R**

171. Rabus, R., Jack, D.L., Kelly, D.J. & Saier, M.H. (1999) *Microbiology* **145**:3431-3445
172. Record, M.T., Reznikoff, W.S., Craig, M.L., McQuade, K.L. & Schlax, P.J. (1996) in: Escherichia coli *and* Salmonella*: cellular and molecular biology*, 2nd edition (eds.: F.C. Neidhardt, R.I. Curtiss, J.L. Ingraham, C.C. Lin, K.B. Low, B. Magasnik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger) ASM Press, Washington D.C., pp. 792-821
173. Reizer, J., Reizer, A. & Saier, M.H. (1997) *Microbiology* **143**:2519-2520
174. Richarme, G. & Dantas Caldas, T. (1997) *J. Biol. Chem.* **272**:15607-15612
175. Rick, P.D. & Silver, R.P. (1996) in: Escherichia coli *and* Salmonella*: cellular and molecular biology*, 2nd edition (eds.: F.C. Neidhardt, R.I. Curtiss, J.L. Ingraham, C.C. Lin, K.B. Low, B. Magasnik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger) ASM Press, Washington D.C., pp. 104-122
176. Riley, M.S., Cooper, V.S., Lenski, R.E., Forney, L.J. & Marsch, T.L. (2001) *Microbiology* **147**:995-1006
177. Rosen, B.P. (1999) *Trends Microbiol.* **7**:207-212
178. Rosenberg, M.F., Callaghan, R., Ford, R.C. & Higgins, C.F. (1997) *J. Biol. Chem.* **272**:10685-10694
179. Rosenberg, M.F., Velarde, G., Ford, R.C., Martin, C., Berridge, G., Kerr, I.D., Callaghan, R., Schmidlin, A., Wooding, C., Linton, K.J. & Higgins, C.F. (2001) *EMBO J.* **20**:5615-5625

**S**


180. Saier, M.H. (2000) *J.Bacteriol.* **182**:5029-5035
181. Saier, M.H. & Reizer, J. (1994) *Mol. Microbiol.* **13**:755-764
182. Salgado, H., Santos-Zavaleta, A., Gama-Castro, S., Millan-Zarate, D., Diaz-Peredo, E., Sanchez-Solano, F., Perez-Rueda, E. & Collado-Vides, J. (2001) *Nucl. Acids Res.* **29**:72-74
183. Sanchez, J.C., Gimenez, R., Schneider, A., Fessner, W.-D., Baldoma, L., Aguilar, J. & Badia, J. (1994) *J. Biol. Chem.* **269**:29665-29669
184. Saraste, M., Sibbald, P.R. & Wittinghofer, A. (1990) *Trends Biochem. Sci.* **15**:430-434
185. Saurin, W. & Dassa, E. (1994) *Prot. Sci.* **3**:325-344
186. Saurin, W., Köster, W. & Dassa, E. (1994) *Mol. Microbiol.* **12**:993-1004
187. Schembri, M.A., Kjærgaard, K. & Klemm, P. (2003) *Mol. Microbiol.* **48**:253-267
188. Schneider, E. & Hunke, S. (1998) *FEMS Microbiol. Rev.* **22**:1-20
189. Shakespeare, W. (1603) *Hamlet, Prince of Denmark*, Stratford upon Avon
190. Shaw, J.G., Hamblin, M.J. & Kelly, D.J. (1991) *Mol. Microbiol.* **5**:3055-3062
191. Simons, R.W., Houman, F. & Kleckner, N. (1987) *Gene* **53**:85-96
192. Simpson, G.L.W. & Ortwerth, B.J. (2000) *Biochim. Biophys. Acta* **1501**:12-24
193. Smith, P.C., Karpowich, N., Millen, L., Moody, J.E., Rosen, J., Thomas, P.J. & Hunt, J.F. (2002) *Molecular Cell* **10**:139-149
194. Sorgen, P.L., Hu, Y., Guan, L., Kaback, H.R. & Girvin, M.E. (2002) *Proc. Natl. Acad. Sci. USA* **99**:14037-14040
195. Sperandio, V., Torres, A.G., Girón, J.A. & Kaper, J.B. (2001) *J. Bacteriol.* **183**:5187-5197
196. Spooner, P.J.R., O'Reilly, W.J., Homans, S.W., Rutherford, N.G., Henderson, P.J.F. & Watts, A. (1998) *Biophys. J.* **75**:2794-2800
197. Stevenson, G., Andrianopoulos, K., Hobbs, M. & Reeves, P. (1996) *J. Bacteriol.* **178**:4885-4893
198. Surette, M.G. & Bassler, B.L. (1998) *Proc. Natl. Acad. Sci. USA* **95**:7046-7050
199. Sutcliffe, I.C. & Russell, R.R.B. (1995) *J. Bacteriol.* **177**:1123-1128


**T**


200. Tam, R. & Saier, M.H. (1993) *Microbiol. Rev.* **57**:320-346
201. Tamames, J. (2001) *Genome Biol.* **2**:1-11
202. Tame, J.R.H., Murshudov, G.N., Dodson, E.J., Neill, T.K., Dodson, G.G., Higgins, C.F. & Wilkinson, A.J. (1994) *Science* **264**:1578-1581
203. Tan, K., Moreno-Hagelsieb, G., Collado-Vides, J. & Stormo, G.D. (2001) *Genome Res.* **11**:566-584
204. Tao, H., Bausch, C., Richmond, C., Blattner, F.R. & Conway, T. (1999) *J. Bacteriol.* **181**:6425-6440
205. Tate, C.G., Muiry, J.A.R. & Henderson, P.J.F. (1992) *J. Biol. Chem.* **267**:6923-6932
206. Tetsch, L. & Kunte, H.J. (2002) *FEMS Microbiol. Lett.* **211**:213-218

207. Tynkkynen, S., Buist, G., Kunji, E., Kok, J., Poolman, B., Venema, G. & Haandrikman, A. (1993) *J. Bacteriol.* **175**:7523-7532

**U**

208. Uebel, S., Kraas, W., Kienle, S., Wiesmüller, K.-H., Jung, G. & Tampé, R. (1997) *Proc. Natl. Acad. Sci. USA* **94**:8976-8981
209. Uebel, S., Plantinga, T., Weber, P.J.A., Beck-Sickinger, A.G. & Tampé, R. (1997) *FEBS Lett.* **416**:359-363
210. Ullmann, R., Gross, R., Simon, J., Unden, G. & Kröger, A. (2000) *J. Bacteriol.* **182**:5757-5764

**V**

211. Van der Does, C., Manting, E.H., Kaufmann, A., Lutz, M. & Driessen, A.J.M. (1998) *Biochemistry* **37**:201-210
212. Van der Heide, T. & Poolman, B. (2000) *Proc. Natl. Acad. Sci. USA* **97**:7102-7106
213. Van der Heide, T. & Poolman, B. (2002) *EMBO Reports* **3**:938-943
214. Van der Sluis, E.O., Nouwen, N. & Driessen, A.J.M. (2002) *FEBS Lett.* **527**:159-165
215. Van Endert, P.M., Saveanu, L., Hewitt, E.W. & Lehner, P.J. (2002) *Trends Biochem. Sci.* **27**:454-461
216. Van Geest, M., Nilsson, I., von Heijne, G. & Lolkema, J.S. (1999) *J. Biol. Chem.* **274**:2816-2823
217. Van Veen, H.W., Margolles, A., Müller, M., Higgins, C.F. & Konings, W.N. (2000) *EMBO J.* **19**:2503-2514
218. Van Veen, H.W., Venema, K., Bolhuis, H., Oussenko, I., Kok, J., Poolman, B., Driessen, A.J.M. & Konings, W.N. (1996) *Proc. Natl. Acad. Sci. USA* **93**:10668-10672
219. Vedder, E., Gossard, S., McCready, M., Ament, J. & Irons, J. (1998) in: *Do the evolution*, on: *Yield*, Sony Music Ltd.
220. Veenhoff, L.M., Heuberger, E.H.M.L. & Poolman, B. (2001) *EMBO J.* **20**:3056-3062
221. Velarde, G., Ford, R.C., Rosenberg, M.F. & Powis, S.J. (2001) *J. Biol. Chem.* **276**:46054-46063
222. Verdon, G., Albers, S.-V., Dijkstra, B.W., Driessen, A.J.M. & Thunissen, A.M. (2003) *J. Mol. Biol.* **330**:343-358
223. Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M. & Lejeune, P. (1998) *J. Bacteriol.* **180**:2442-2449
224. Vigano, C., Grimard, V., Margolles, A., Goormaghtigh, E., Van Veen, H.W., Konings, W.N. & Ruysschaert, J.-M. (2002) *FEBS Lett.* **530**:197-203
225. Vigano, C., Margolles, A., Van Veen, H.W., Konings, W.N. & Ruysschaert, J.-M. (2000) *J. Biol. Chem.* **275**:10962-10967
226. Von Heijne, G. (1986) *EMBO J.* **5**:3021-3027
227. Von Heijne, G. (1992) *J. Mol. Biol.* **225**:487-494
228. Vos, J.C., Spee, P., Momburg, F. & Neefjes, J. (1999) *J. Immunol.* **163**:6679-6685

**W**


229 Wachowski, A., and Wachowski, L. (1999) in: *The Matrix*, Warner Bros., Los Angeles

230. Walker, J.E., Saraste, M., Runswick, M.J. & Gay, N.J. (1982) *EMBO J.* **1**:945-951

231. Walmsley, A.R., Shaw, J.G. & Kelly, D.J. (1992) *Biochemistry* **31**:11175-11181

232. Walmsley, A.R., Shaw, J.G. & Kelly, D.J. (1992) *J. Biol. Chem.* **267**:8064-8072

233. Wei, Y., Lee, J.-M., Richmond, C., Blattner, F.R., Rafalski, J.A. & LaRossa, R.A. (2001) *J. Bacteriol.* **183**:545-556

234. Wilde, O. (1891) in: *Preface to 'The picture of Dorian Grey'*, Penguin Books Ltd., London

235. Williams, K.A. (2000) *Nature* **403**:112-115

236. Williams, K.A., Geldmacher-Kaufer, U., Padan, E., Schuldiner, S. & Kühlbrandt, W. (1999) *EMBO J.* **18**:3558-3563

237. Wolf, A., Shaw, E., Nikaido, K. & Ames, G.F.-L. (1994) *J. Biol. Chem.* **269**:23051-23058

238. Wolf, A., Shaw, E., Oh, B.-H., De Bondt, H., Joshi, A.K. & Ames, G.F.-L. (1995) *J. Biol. Chem.* **270**:16097-16106

239. Wolf, Y.I., Rogozin, I.B., Kondrashov, A.S. & Koonin, E.V. (2001) *Genome Res.* **11**:356-372

240. Wood, N.J., Alizadeh, T., Richardson, D.J., Ferguson, S.J. & Moir, J.W.B. (2002) *Mol. Microbiol.* **44**:157-170

241. Wyborn, N.R., Alderson, J., Andrews, S.C. & Kelly, D.J. (2001) *FEMS Microbiol. Lett.* **194**:13-17


**X**


242. Xavier, K.B., Martins, L.O., Peist, R., Kossmann, M., Boos, W. & Santos, H. (1996) *J. Bacteriol.* **178**:4773-4777


**Y**


243. Yasueda, H., Kawahara, Y. & Sugimoto, S.-I. (1999) *J. Bacteriol.* **181**:7154-7160

244. Yew, W.S. & Gerlt, J.A. (2002) *J. Bacteriol.* **184**:302-306

245. Yorke, T., Selway, P., O'Brien, E., Greenwood, C. & Greenwood, J. (2003) in: *Mimminent*, on: *Hail to the Thief - special edition*, EMI Records Ltd.

246. Yura, T. & Nagai, H. (1993) *Annu. Rev. Microbiol.* **47**:321-350


**Z**


247. Zhang, R., Kim, Y., Skarina, T., Beasley, S., Laskowski, R., Arrowsmith, C., Edwards, A., Joachimiak, A. & Savchenko, A. (2002) *J. Biol. Chem.* **277**:19183-19190

248. Zhang, Z., Aboulwafa, M., Smith, M.H. & Saier, M.H. (2003) *J. Bacteriol.* **185**:2243-2250

# *Appendix: Theses*

1. A good these not only is of current value, but is also ageless.

2. Knowledge of the transported substrate is essential for researching the activity of a transport protein. (*this thesis*)

3. The unpublished chapters of scientific theses are of great importance to fellow scientists.

4. Even inexplicable results can be reproducable. (*this thesis*)

5. Dogmas are a death-blow to scientific research.

6. A good teacher forever remains pupil.

7. The creation of more jobs with perspective at universities will lead to a reduction in the number of scientists that pursue a career outside of science.

8. The ability to critically listen to music is less alcohol-sensitive than the ability to produce beautiful music.

9. Due to the frequent recycling of existing melodies, modern day pop music can be considered as 'not harmful to the environment'.

10. The appearance of a mussel is not compensated by its taste.

11. The relief of discovering an Ecuadorian public toilet disappears as soon as its door has been opened.

12. The phrase 'knowledge is power' for many only becomes true after finishing secondary school.

13. 'Independent' does not mean 'alone'.

14. Not everything tastes better with pineapple. (*in response to the thesis of C. van der Does*)